# Dynamic Persuasion and Strategic Search

Yunfei (Jesse) Yao*

*(University of California, Berkeley)*

May 16, 2022

Latest Version: *Click Here*

* Comments welcome. I thank J. Miguel Villas-Boas, Ganesh Iyer, Yuichiro Kamada, Shachar Kariv, Ryota Lijima, Qitian Ren, Chris Shannon, Dong Wei, Pinar Yildirim, and seminar participants at Berkeley Haas and Berkeley Econ theory lunch for constructive comments. All errors are my own. E-mail address: jesseyao@berkeley.edu.

# Abstract

We consider a dynamic model in which a sender sequentially persuades a receiver to take the desired action regardless of the unknown state. In contrast, the receiver only wishes to take that action if the state is good. The receiver could incur costs to search for more information. The updated belief about the state helps him[1] make decisions. The sender controls the information environment. Given the information provision strategy of the sender, the receiver trades off the benefit and cost of information acquisition and decides whether to search or not. The sender could incur higher costs to provide more information so that the receiver would be more likely to take the desired action. Given the information acquisition strategy of the receiver, the sender trades off the benefit and cost of information provision and decides how much information to provide. She smoothens the information provision over multiple periods if and only if the prior of the state is good is high. The optimal information structure fully reveals the state being good when the search cost is high and partially when the search cost is low. We also compare the information structures between the profit-maximizing one and the efficient one. The profit-maximizing strategy is efficient if and only if the search cost is high.


[1] We refer to the sender as “she” and the receiver as “he” throughout the paper.

# 1 Introduction

With the rapid development of technologies, it is more and more common for an individual to acquire some information before making a decision. More information could lead to less uncertainty and help the decision-making. As the learning outcome depends crucially on the information environment, other related parties have strong incentives to influence it so that the learning outcome would be more favorable to them. For example, a consumer considering purchasing a product may search on Google to find out whether the product matches his need or not. The seller could control the type of information available to the consumer by search advertising. Furthermore, she may provide information more than once through different channels by retargeting. By providing different kinds of information, the seller affects the search and purchasing decisions of the consumer. The costs of the seller depend on the information structure. If she wants the potential buyer to see the information about the product more often on Google, she needs to spend more money to bid for the search advertising spots. The FDA needs to decide whether to approve or disapprove many new drugs every year. The pharmaceutical company may design tests to convince the regulator that the drug is safe. Monitoring and examining the test is costly for the regulator. So, the regulator would only examine the test if the likelihood that the result is positive and the precision of the test is high enough. In the meantime, different tests cost differently for the company. Even if a test fails to convince the regulator, the firm could offer another test, hoping that the regulator would approve the drug after seeing the new test.

To better understand the above scenarios, we consider a dynamic model in which a sender sequentially persuades a receiver to take the desired action (denote it by action $G$) regardless of the unknown state. In contrast, the receiver only wishes to take that action if the state is good. The receiver could incur costs to search for more information. The updated belief about the state helps him make decisions. If the receiver observes bad news, he knows that the state is less likely to be good and will not take action $G$ without the arrival of new information. If the receiver observes good news, he knows that the state is more likely to be good, and the expected payoff of taking action $G$ increases. The sender controls the information environment. Given the information provision strategy of the sender, the receiver trades off the benefit and cost of information acquisition and decides whether to search or not. The receiver is forward-looking and forms rational expectations of the sender's strategy. The sender could incur higher costs to provide more information so that the receiver would be more likely to take action $G$. Given the information acquisition strategy of the receiver, the sender trades off the benefit and cost of information provision and decides how much information to provide. Therefore, the receiver and the sender are trading off the benefit and cost of information acquisition/provision simultaneously. At each period, the sender chooses the information structure, and the receiver chooses whether to search for more information or make a decision. This paper takes into account the above features and characterizes the optimal information provision strategy of the sender and the optimal information acquisition strategy of the receiver. In equilibrium, the sender uses one-shot experiments, which induce the receiver to take action $G$ immediately upon observing good news. This way, the sender saves the expected search time and does not need to compensate the receiver for a higher expected search cost. The sender smoothens the information provision over multiple periods if and only if the prior of the state is good is high. The sender extracts all the surplus from the receiver when she provides information in both periods while may leave some rent to the receiver when she provides information in only one period. As the information provision cost is convex, the sender has an incentive to smoothen the information provision over two periods. However, the higher expected search cost of the receiver would discourage him from searching if the likelihood of getting a strictly positive ex-post payoff (taking action $G$ at state $g$) is low. Hence, the sender only smoothens the information provision over multiple periods if the prior is high. The optimal information structure fully reveals the state being good when the search cost is high and partially when the search cost is low. The above result comes from the tradeoff between the frequency and precision of the signal.

We also compare the information structures between the payoff-maximizing one and the efficient one. When the search cost is high, there is no information distortion, and the optimal strategy of the sender is efficient. There could be either upward or downward information distortion when the search cost is lower.

In the main model, the sender can commit to the current-period information structure but cannot promise to the receiver the information structure in the future. In the extension, we study the implications of dynamic commitment power. We find that the ability to commit to the strategy in the future strictly benefits the sender when the search cost is high, while the benefit vanishes as the search cost approaches zero. We also consider the implications of discounting. When the search cost is high, the participation constraints of the receiver are hard to satisfy. Providing enough information to persuade the receiver to search dominates the force of discounting. So, the optimal strategy of the sender does not depend on the discount factor. When the search cost is low and the players become less patient, the present value of the second-period profit decreases and the first-period participation constraint would be tighter. The sender has a stronger incentive to convert the receiver early. So, she provides (weakly) less information in the second period and (weakly) more information in the first period. Lastly, we consider a myopic receiver who only cares about the current-period payoff. We find that both payers are (weakly) worse off if the receiver is myopic rather than forward-looking. The managerial implication is that the sender should try to better inform the receiver of the possibility of multi-period information revelation and gradual learning.

## 1.1 Related Literature

There is a large stream of literature on optimal information acquisition. Branco et al. (2012), Ke and Villas-Boas (2019), and Ke et al. (2016) study the optimal search strategy of the consumer in which the decision maker decides when to stop searching. Moscarini and Smith (2001) studies the optimal information acquisition strategy when consumers could learn at a higher rate by incurring a higher flow cost. In the above papers, the information environment is exogenous. As the strategy and the outcome depend on the information structure, other parties (e.g., the seller) have strong incentives to influence it. In our paper, we take it into account by having the sender determine the information environment of the receiver. In related literature on choice overload, the sender determines the amount of information provided to the receiver (Branco et al. 2016, Kuksov and Villas-Boas 2010). The amount of information provided by the sender affects the sender's payoff through its impact on the receiver's search/evaluation cost. Only the receiver trades off the benefit and cost of search in those papers. In our model, we consider the cost for the sender to provide information as well.

We use a belief-based method of modeling information provision, first introduced by Aumann and Maschler (1995) and Kamenica and Gentzkow (2011) in the Bayesian persuasion and information design literature. The sender picks a mean-preserving spread of the prior as the posterior belief, which simplifies the analyses. Gentzkow and Kamenica (2014) extends the concavification approach of Kamenica and Gentzkow (2011) to the setting where the sender's cost depends on the expected reduction in uncertainty of the signals. Degan and Li (2021) studies the persuasion problem in which the persuasion cost depends on the precision of the signals. Wei (2021) considers the receiver's cost of learning by introducing a rationally inattentive receiver who could process less information than what the sender provides. Subsequent papers have extended the static setting to the dynamic one, but it is usually costless for the sender to provide information (Ely 2017, Ely and Szydlowski 2020, Iyer and Zhong 2021, Orlov et al. 2020). In Che et al. (2020), the sender performs costly experimentation to persuade the receiver to take a particular action. The flow cost of providing the information is constant, and the type of experimentation available to the sender is limited. In our paper, the sender could speed up the receiver's learning by incurring a higher cost. So, we could investigate the optimal tradeoff of the benefit and cost of information provision by the sender. Li and Norman (2021) and Wu (2021) study a different type of sequential Bayesian persuasion problem. Multiple senders communicate to the receiver sequentially, each sending a single signal. In this paper, the same sender sends multiple signals to the receiver sequentially. In addition, unlike their papers, the receiver voluntarily listens to the sender in our model. So, the sender takes into account the participation constraints in designing the experiments.

The remainder of the paper is organized as follows. Section 2 presents the main model. Section 3 characterizes the optimal information provision strategy and the equilibrium outcomes. Section 4 characterizes the efficient information provision strategy and summarizes the information distortion when the sender rather than the social planner designs the information structure. Section 5 includes several model extensions. Section 6 concludes.

# 2 The Model

## 2.1 States, Actions, and Payoffs

There are two players, a sender and a receiver, and two states, good ($g$) and bad ($b$). The receiver ultimately makes a binary decision between $G$ and $B$. The sender wishes to persuade the receiver to take action $G$ regardless of the state while the receiver wishes to match the decision with the state (taking action $G(B)$ when the state is $g(b)$). The payoffs of the decision for the players are the following:

| (sender payoff, receiver payoff) | action G | action B |
|---|---|---|
| state $g$ | $(p, v_g)$ | $(0, 0)$ |
| state $b$ | $(p, v_b)$ | $(0, 0)$ |

We assume that the sender's payoff when the receiver takes the desired action is positive, $p > 0$. The receiver's payoff is positive if he takes action $G$ when the state is $g$, $v_g > 0$, and negative if he takes action $G$ when the state is $b$, $v_b < 0$. Both players get zero payoff if the receiver takes the action $B$ (which could be thought of as a quitting option). We further assume without loss of generality that $v_g = 1 + v_b$.[2] Both the sender and the receiver do not know the state initially but have a common prior belief on it, $\mu_0 := \mathbb{P}(\text{the state is } g) \in (0, 1)$. Each period $t \in \{0, 1\}$, the sender determines and commits to the information structure of the experiment but cannot commit to the experiment in the future. The receiver could search

[2] To see that assuming $v_g = 1 + v_b$ is WLOG, consider the following normalization. Let $v_g' = \frac{v_g}{v_g - v_b}, v_b' = \frac{v_b}{v_g - v_b}, p' = \frac{p}{v_g - v_b}$. Then, $v_g' - v_b' = 1$. We make this assumption to simplifiy the analyses and the presentation.

for information (action $S$) before deciding. The information acquisition is costly but helps the receiver to make better decisions. If the receiver chooses to search, he would observe an experiment that reveals some information about the state. The players update the belief after the realization of the experiment.[3] The game ends whenever the receiver makes a decision ($G$ or $B$). Figure 1 illustrates the timing of the game.

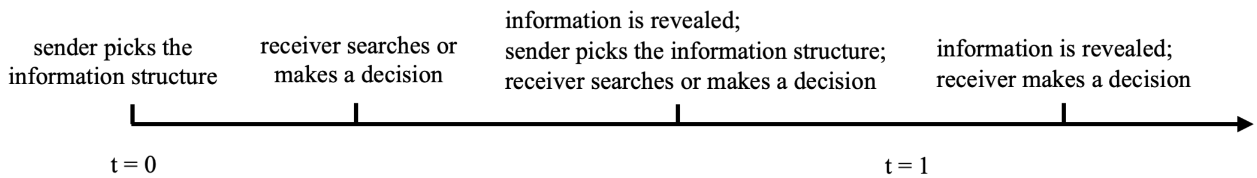


Figure 1: Timing of the Game

We model the experiment by a binary signal $s \in \{0, 1\}$. $\mathbb{P}[s = 1|g]$ and $\mathbb{P}[s = 1|b]$ uniquely determines the signal. We order the value of the signal such that $\mathbb{P}[s = 1|g] > \mathbb{P}[s = 1|b]$. Hence, $s = 1$ corresponds to good news and $s = 0$ corresponds to bad news. Analogous to Proposition 1 of Kamenica and Gentzkow (2011), we could work with mean-preserving posterior beliefs rather than the specific signal structure to simplify the analyses. Specifically, the existence of a binary signal $s$ such that $\mathbb{P}[s = 1|g] > \mathbb{P}[s = 1|b]$ is equivalent to the existence of a binary-valued posterior belief whose expectation is equal to the prior.[4]

Denote the belief at the beginning of each period by $\mu_t$. At each period, with probability $\lambda_t$, the receiver observes good news, and the belief increases to $\bar{\mu}_t$. We refer to $\lambda_t$ as the arrival rate of good news and $\bar{\mu}_t$ as the belief after observing good news. With probability $1 - \lambda_t$, the receiver observes bad news, and the belief decreases to $\underline{\mu}_t$. We refer to $\underline{\mu}_t$ as the belief after observing bad news.

We assume that there is no discounting in the main model, and discuss the implication of discounting in the extension. The receiver incurs a flow cost of $c$ per period of search. The sender's cost of information provision depends on the arrival rate of good news, $K = K(\lambda)$. The sender incurs a higher cost to provide information structures with more frequent good news. Throughout this paper, we refer to the arrival rate of good news as the amount of information. More information means more frequent good news. We make the following assumptions on the cost of the sender.

[3] We can assume WLOG that the sender observes the realization of the signal because her optimal payoff without observing the realization of the signal is bounded from above by the optimal payoff if she could observe it. We will show in Proposition 1 that the sender could perfectly infer the realization of the signal from the receiver's action under the optimal signal structure of the latter case. So, the equilibrium outcomes with and without observing the realization of the signal coincide.

[4] In the appendix, we state this result formally and provide proof.

**Assumption 1.** $K(\cdot) \in \mathcal{C}^2(\mathbb{R}_+), K'(\lambda) > 0, K''(\lambda) > 0,\ K(0) = 0,\ \lim_{\lambda \to 1^-} K'(\lambda) = +\infty$, $\lim_{\lambda \to 0^+} K'(\lambda) = 0$.

This assumption means that it is relatively cheap for the sender to provide a little information. The marginal cost of the sender increases at an increasing rate when the amount of information provided increases.

The total payoff for each player is the payoff of the decision net of the information provision/acquisition costs. The receiver is forward-looking and forms rational expectations about the sender's strategy in the future. We assume that $\mu_0 v_g + (1-\mu_0)v_b < 0 \Leftrightarrow \mu_0 < -v_b$. So, the receiver would never take action $G$ without search. Otherwise, the sender would simply provide no information and the receiver always takes the desired action. We also assume that the search cost is not too high, $c < v_g$. Otherwise, the receiver would never search.

## 2.2 Applications

### 2.2.1 Product Sales

A seller offers a product with unit demand and 0 marginal costs. The price is $p$. The product may be a good match or a bad match with the buyer. Both the buyer and the seller do not know the state initially but have a common prior belief. The product is of zero value to the buyer if the state is bad and is of value $v$[5] to the buyer if the state is good. The buyer decides on purchasing the product or not. The payoffs of the decision for the players are the following:

| (sender payoff, receiver payoff) | purchasing | not purchasing |
|---|---|---|
| state good match | $(p, v-p)$ | $(0, 0)$ |
| state bad match | $(p, -p)$ | $(0, 0)$ |

The buyer could incur costs to search for more information about the product. The updated belief helps him to make better purchasing decisions. If the buyer observes bad news, he knows that the product is less likely to be a good match and avoids wasting money on it without the arrival of new information. If the buyer observes positive news, he knows that the product is more likely to be a good match, and the expected valuation increases.

[5] By normalizing the payoffs, we can set $v = 1$ WLOG.

The buyer gets a positive surplus upon purchasing the good if the expected value is higher than the price.

### 2.2.2 New Drug Launches

A pharmaceutical company develops a new drug and tries to get approval from the regulator (e.g., the FDA) by designing and conducting tests to convince the regulator that the drug is safe. If the regulator approves it, the firm gains an expected profit of $p$. The regulator gains a positive utility if the drug is safe and a negative utility if it is unsafe. The payoffs of the decision for the players are the following:

| (sender payoff, receiver payoff) | approval | disapproval |
|---|---|---|
| state safe | $(p, v_g > 0)$ | $(0, 0)$ |
| state unsafe | $(p, v_b < 0)$ | $(0, 0)$ |

Monitoring and examining the test is costly for the regulator. So, the regulator would only investigate the test if both the likelihood that the result is positive and the precision of the test are high enough. Even if a single test fails to convince the regulator, the firm could offer another test, hoping that the result would finally convince the regulator to approve the drug.

## 2.3 Strategies and Equilibrium Concepts

Since the belief is common knowledge and there is no private information, we consider the sender-preferred subgame perfect equilibrium, as in Kamenica and Gentzkow (2011). If multiple actions ($G$, $B$, and $S$) give the receiver the same expected payoff, we assume that the receiver chooses an action that maximizes the sender's expected payoff. We also assume that the sender prefers to give the receiver more surplus in the first period if more than one equilibria lead to the same expected sender's payoff. A perturbation of very little discounting justifies this assumption.

### 2.3.1 Equilibrium of the Subgame

In the second period, the sender and the receiver play the subgame. Since it is costly for the sender to provide information, the sender would either give no information or provide enough information succh that the receiver searches and would take action $G$ if good news arrives. We illustrate the belief evolution of the subgame in Figure 2. Also, the distribution of the belief induced by the experiment should be a mean-preserving spread of the initial belief: $\mathbb{E}[\Delta\mu] = 0$. In sum, the sender either does not provide information and obtains zero payoffs or takes into account the following constraints when designing the information structure:

(1) participation constraint:

$$\lambda_1[\bar{\mu_1} v_g + (1-\bar{\mu_1})v_b] = \lambda_1(\bar{\mu_1} + v_b) \geq c \tag{$IR_1$}$$

(2) feasibility constraint:

$$\lambda_1\bar{\mu_1} + (1-\lambda_1)\underline{\mu}_1 = \mu_1 \tag{$F_1$}$$

If the sender provides information, the constrained program of the sender is:

$$\max_{\lambda_1,\bar{\mu_1}} -K(\lambda_1) + p\lambda_1 \tag{$P_1$}$$
$$\text{s.t. } (IR_1), (F_1), \lambda_1 \in [0,1], \underline{\mu}_1 \in [0, \mu_1)$$

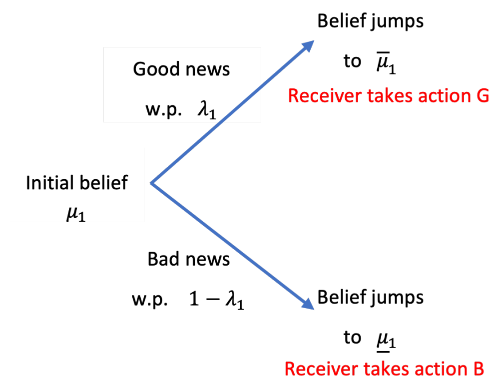


Figure 2: Belief Evolution of the Subgame

We analyze the solution to this problem in the next section. Though the information structure consists of $(\lambda_1, \bar{\mu_1}, \underline{\mu}_1)$, any two of them fully characterize the strategy as the third variable is then uniquely determined by $(F_1)$. Therefore, we use $(\lambda_1, \bar{\mu_1})$, the arrival rate of good news and the belief after observing good news, to represent the sender's strategy of the subgame.

### 2.3.2 Equilibrium of the Full Game

The sender has three options. Firstly, she could provide no information and obtain zero payoffs. Secondly, she could provide information in only one period. If the receiver decides to search, he would observe a one-shot experiment designed by the sender. The receiver takes action $G$ if good news arrives and takes action $B$ if bad news comes. The sender would not provide extra information regardless of the experimental outcome. Her problem is exactly ($P_1$) (except that the subscript 1 shall be replaced by 0). Lastly, the sender could provide information in both periods. She could give a pair of one-shot experiments, which means that the receiver would take action $G$ upon observing good news at either period. If bad news arrives in the first period, the sender will provide another experiment, hoping that good news will come in the second period. The sender could also offer a pair of iterative experiments. The receiver takes action $G$ if and only if he observes good news at both periods. We illustrate the belief evolution of the one-shot experiments and iterative experiments in Figure 3 and 4.

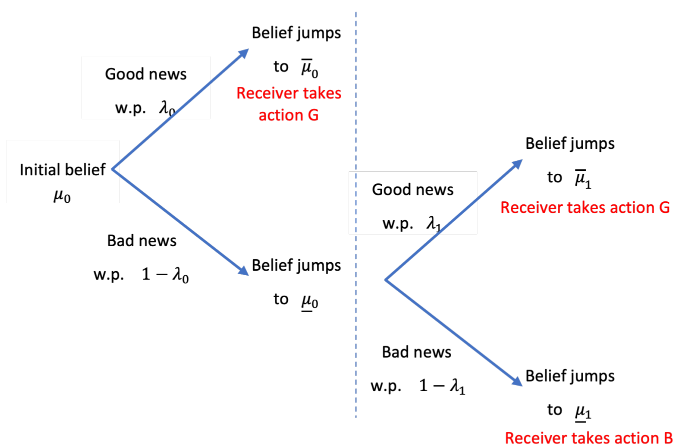


Figure 3: Belief Evolution of the One-shot Experiments

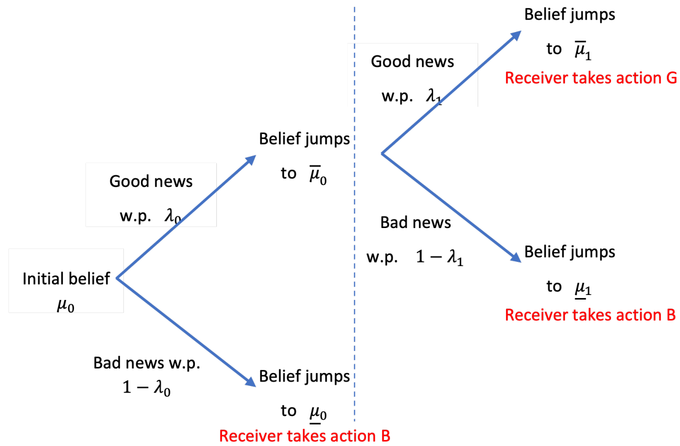


Figure 4: Belief Evolution of the Iterative Experiments

The receiver must search in both periods before taking action $G$ under iterative experiments. Hence, compared to one-shot experiments, iterative experiments requires a longer search time. To compensate the receiver for the higher expected search costs, the sender needs to provide information more favorable to the receiver, which hurts the sender's payoff. The following result shows that the sender always prefers one-shot experiments in equilibrium. So, we limit our attention to the optimal one-shot experiments.

**Proposition 1.** *For any pair of feasible iterative experiments, there exists a one-shot experiment that gives the sender a strictly higher payoff.*

The sender takes into account the following constraints when designing the optimal one-shot experiment of the first period:
(1) participation constraint:

$$\begin{aligned}&\lambda_0[\bar{\mu_0}v_g + (1-\bar{\mu_0})v_b] + (1-\lambda_0)\mathbb{E}[\text{receiver surplus at } t=1|\text{search at } t=1]\\ =&\lambda_0(\bar{\mu_0}+v_b) + (1-\lambda_0)[\lambda_1(\bar{\mu_1}+v_b)-c] \geq c \end{aligned} \tag{$IR_0$}$$

(2) feasibility constraint:

$$\lambda_0\bar{\mu_0} + (1-\lambda_0)\underline{\mu}_0 = \mu_0 \tag{$F_0$}$$

If the sender provides information in both periods, her problem is:[6]

[6] We impose the following implicit assumptions on all of the programs in the remaining of the paper: $\bar{\mu_t} \in [-v_b, 1], \underline{\mu_t} \in [0, \mu_t), \lambda_t \in [0,1], \underline{\mu}_0 = \mu_1$. The last equality comes from the fact that the belief at the beginning of the second period, $\mu_1$, is the belief after observing bad news in the first period, $\underline{\mu}_0$, under one-shot experiments.

$$\max_{\lambda_0,\bar{\mu_0},\mu_1,\lambda_1,\bar{\mu_1}} -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\lambda_1)+p\lambda_1\right] \tag{$P_2$}$$
$$\text{s.t. } (IR_0),(F_0),(\lambda_1,\bar{\mu}_1) \text{ solves } (P_1)$$

We analyze the solution to this problem in the next section. We use $(\lambda_0, \bar{\mu_0}, \mu_1, \lambda_1, \bar{\mu_1})$, the arrival rate of good news at each period, the belief after observing good news at each period, and the initial belief in the second period, to represent the sender's strategy of the entire game.

# 3 Optimal Strategies

From the previous discussion, the receiver would search (take action $S$) whenever the sender provides information. The receiver would take action $G$ immediately upon receiving good news and $B$ if the sender does not provide information. Therefore, we only need to characterize the sender's strategy, which implies the receiver's strategy.

## 3.1 A Benchmark

We first consider a benchmark problem in which the receiver always participates, and the sender could generate an arbitrary amount of information at each period. The sender chooses the information structure to maximize the expected payoff. We will use the solution throughout the subsequent analyses.

$$\max_{\lambda_0,\lambda_1} -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\lambda_1)+p\lambda_1\right] \tag{$P_b$}$$

**Lemma 1.** *The solution to the benchmark problem $(P_b)$, $(\lambda_0^{**}, \lambda_1^{**})$, does not depend on the search cost $c$ and $\lambda_0^{**} < \lambda_1^{**}$. The benchmark sender's payoff is strictly positive.*

The benchmark sender's payoff is the highest possible payoff the sender could obtain in equilibrium. When the prior is high enough, good news arrives at the rate of $\lambda_t^{**}$ and the receiver searches. The sender obtains the benchmark payoff. When the prior is lower and good news arrives at the benchmark arrival rate, the receiver may be quite uncertain about the state even after observing good news (low $\bar{\mu}_t$) due to feasibility constraints. So, the receiver chooses not to search. This friction restricts the type of communication the sender and the receiver could have and distorts the optimal strategy. For the problem to be non-degenerate, we concentrate on the case in which the prior is not too high such that the optimal strategy is different from the benchmark solution throughout the paper.

## 3.2 Optimal Strategy of the Subgame

When the belief at the beginning of the second period is too low, the receiver will not search given any feasible experiments. Thus, the sender does not provide information to minimize the cost. When the belief at the beginning of the second period is higher, and the search cost is not too high, the sender provides information and obtains a positive payoff. The following proposition summarizes the optimal information structure.

**Proposition 2.** *At the second period, the sender does not provide information when $\mu_1 < \mu_{0,1} := c/v_g$. When $\mu_1 \geq \mu_{0,1}$, the optimal arrival rate of good news and the optimal belief after observing good news, $(\lambda_1^*, \bar{\mu}_1^*)$, depend on the search cost $c$:*

1. *If $c \geq v_g\lambda_1^{**}$, there exists a unique $\widehat{c} \in (v_g\lambda_1^{**}, v_g\mu_1]$ such that the sender does not provide information if $c > \widehat{c}$ and $(\lambda_1^*, \bar{\mu}_1^*) = (c/v_g, 1)$ if $c < \widehat{c}$. The receiver gets zero surplus.*

2. *If $c \in [\mu_1 + v_b\lambda_1^{**}, v_g\lambda_1^{**})$, $(\lambda_1^*, \bar{\mu}_1^*) = (\frac{\mu_1 - c}{-v_b}, \frac{-v_b\mu_1}{\mu_1 - c})$. The receiver gets zero surplus.*

3. *If $c < \mu_1 + v_b\lambda_1^{**} \wedge v_g\lambda_1^{**}$, $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}} \wedge 1)$. The receiver gets strictly positive surplus.*

When the search cost is too high, the sender has to provide a lot of information to persuade the receiver to search. Even if it is feasible for the sender to provide enough information that the receiver would search, it is so costly that the expected sender's payoff is negative. So, the sender chooses not to provide information, and the receiver does not search.

When the search cost is high but not too high, the sender will provide just enough information such that the receiver searches. Since the receiver's participation constraint is hard to satisfy, in equilibrium, the receiver becomes certain ($\bar{\mu}_1 = 1$) that the state is $g$ after observing good news. Suppose, instead, a good signal does not fully reveal the state ($\bar{\mu}_1 < 1$), its arrival rate needs to be higher to persuade the receiver to search. Since the marginal cost of increasing the arrival rate of good news exceeds the marginal benefit, the sender's payoff decreases. The sender trades off the frequency of good news for precision.

When the search cost is intermediate, the receiver's participation constraint is easier to satisfy. Since the marginal benefit of increasing the arrival rate of good news exceeds the marginal cost, in equilibrium, the sender trades off the precision of good news for the frequency. The receiver is still uncertain about the state after observing good news. But the belief would be high enough such that the receiver searches.

When the search cost is low, the information friction does not distort the information structure. The sender provides the benchmark amount of information, and the receiver gets a strictly positive surplus.

## 3.3 Optimal Strategy of the Entire Game

When the prior is too low, any feasible signal the sender could generate is not attractive enough for the receiver to search. Thus, it is impossible to communicate between the sender and the receiver. When the prior is higher, and the search cost is not too high, the sender provides information and obtains a positive payoff.

**Proposition 3.** *Suppose the search cost is not too high, $c < \widehat{c}$. There exists $\mu_{1,2} \geq c(2v_g - c)/(v_g)^2$ such that the sender does not provide information if the prior is low, $\mu_0 < \mu_{0,1}$, provides information in one period if $\mu_0 \in [\mu_{0,1}, \mu_{1,2})$, and provides information in both periods if $\mu_0 > \mu_{1,2}$. Suppose the sender provides information in both periods. Good news fully reveals the state (being $g$) when the search cost is high and partially reveals the state when the search cost is low. The receiver gets zero total surplus.*

Since the information provision cost is non-linear in the arrival rate of good news, the sender has an incentive to smoothen the information provision over two periods. When the prior is low, it is not feasible for the sender to provide enough information in both periods so that the receiver would search whenever good news has not arrived. As the prior increases, it becomes feasible for the sender to smoothen the information provision. If the sender finds it optimal to provide information in both periods at a given prior, she also prefers smoothening information for any higher prior.

When it is highly costly for the receiver to search, the optimal information structure fully convinces the receiver that the state is $g$ when good news arrives. In equilibrium, the receiver obtains the highest possible surplus conditional on observing good news and making the purchase. Without providing this type of information favorable to the receiver, the sender cannot persuade the receiver to search. In contrast, when it is less costly for the receiver to search, the optimal information structure adds some noise to good news. In equilibrium, the receiver is not sure that the state is $g$ after observing good news. The state may be $b$ after the receiver takes action $G$. However, the likelihood of state $g$ after good news is high enough to persuade the receiver to search. By adding some noise to good news, the sender could provide more frequent good news and increase her payoff without violating the feasibility constraint.

When the sender provides information in both periods, she can always extract surplus from the receiver and increases her payoff if the receiver gets a strictly positive surplus. If the sender faces an information under-provision issue, she can increase the payoff by increasing the arrival rate of good news and decreasing the belief after observing good news. If the sender faces an information over-provision issue, she can increase the payoff by reducing the arrival rate of good news and increasing the belief after observing good news. It implies that the receiver gets zero surpluses in equilibrium.

The optimal strategy is consistent with real-world examples. Consider the application of product sales in section 2.2.1 when a buyer decides whether to purchase a pair of shoes and searches on Google to gather some information. The search cost is relatively low, and the buyer is less likely to be sure that it is a good match even after obtaining good news and making the purchase. On the contrary, when a buyer visits a dealership and acquires information on the car of interest through talking to the dealer and test driving it, the search cost is relatively high as it is costly to travel to the physical store and spend a significant amount of time there. When the buyer decides to buy the car, he is usually confident that it is precisely a good match.

## 3.4 Comparative Statics

When the sender provides information in only one period, the optimal strategy has a closed-form solution and is easy to analyze. Here, we discuss the comparative statics when the sender provides information in both periods. The specific form of the optimal strategy depends on the relative size of the search cost. According to Proposition 2, the sender's strategy of the subgame is constant when the search cost is high (but not too high). When the search cost is lower, it depends on the belief at the beginning of the second period, $\mu_1$. When $\mu_1 > c - v_b\lambda_1^{**}$, the receiver expects to get a strictly positive surplus in the second period, and thus the first-period participation constraint is relaxed (denote the corresponding strategy by $S_+$ strategy). When $\mu_1 \in [c/v_g, c - v_b\lambda_1^{**}]$, the receiver expects to get zero surpluses in the second period (denote the corresponding strategy by $S_0$ strategy). In equilibrium, the sender endogenously determines whether to use the $S_0$ or $S_+$ strategy.

### 3.4.1 Comparative Statics about the Prior

When the search cost is high, the prior determines whether the sender smoothens the information but does not affect the information structure conditional on the sender providing information. When the search cost is low, the prior affects the information structure monotonically. When the search cost is intermediate, the sender may switch from the $S_0$

strategy to the $S_+$ strategy as the prior increases. There could be a discrete jump in the optimal information structure. We leave the analysis of this case in the appendix.

**Proposition 4.** *Suppose the sender provides information in both periods. When the search cost is high, $v_g\lambda_1^{**} \leq c < \widehat{c}$, good news fully reveals the state. The arrival rate of good news, $\lambda_t^*$, and the sender's payoff does not depend on the prior, $(\lambda_t^*, \bar{\mu}_t{}^*) = (c/v_g, 1)$. When the search cost is low, $c \leq \tilde{c} := v_g K'^{-1}\left[K(\lambda_1^{**})/\lambda_1^{**}\right]$, the arrival rate of good news, $\lambda_t^*$, is continuous and increasing in the prior. The belief after observing good news, $\bar{\mu}_t^*$, is continuous and decreasing in the prior. The sender's payoff is strictly increasing in the prior.*

The optimal information is perfectly smooth when the search cost is high, and the sender provides information in both periods. Since the participation constraint of the receiver is strong, good news fully reveals the state is $g$. As it is very costly for the receiver to acquire information, the minimal amount of information to persuade the receiver to search is high. The marginal cost of providing more information exceeds the marginal benefit. Even if the prior increases and it is feasible for the sender to provide more information, she would prefer not to do so. Hence, conditional on the sender providing information, the information structure does not depend on the prior.

When the search cost is low, and the sender provides information in both periods, she chooses between $S_+$ and $S_0$ strategies. Under the $S_+$ strategy, the receiver observes less frequent good news in the first period and more frequent good news in the second period. On average, he spends a longer time searching. Consequently, the sender has to provide information more favorable to the receiver to compensate for the higher expected total search cost, which hurts the sender's surplus. Therefore, the sender always chooses the $S_0$ strategy in equilibrium, and the optimal strategy is continuous in the prior. The sender faces an information under-provision issue in both periods. More frequent good news is feasible when the prior is higher. Even if the receiver becomes less sure about the state being good after observing good news, he would still search as long as the likelihood of receiving good news and earning a strictly positive surplus increases. In equilibrium, the sender trades off the precision of good news for the frequency as the prior increases.

### 3.4.2 Comparative Statics about sender's Costs

Providing the same amount of information may cost the sender differently. To study the impact of sender's information provision costs on the optimal strategy, we rewritten the sender's cost function by $K(\lambda) = \eta\tilde{K}(\lambda)$, where $\tilde{K}(1/2) = 1$ for identification. It is more costly for the sender to provide information when $\eta$ is larger. The following proposition summarizes the comparative statics of the optimal strategy about $\eta$.

**Proposition 5.** *Suppose the sender provides information in both periods. Her payoff is strictly decreasing in the relative cost of information provision, $\eta$. When the search cost is low, $c \leq \tilde{c}$, the sender provides (weakly) less information in the first period and (weakly) more information in the second period, as $\eta$ increases. When the search cost is high, $c \geq v_g \lambda_1^{**}$, the optimal strategy of the sender does not depend on $\eta$.*

When the search cost is high, it is very costly for the receiver to search. The sender needs to provide a lot of information to persuade the receiver to search. As a result, the marginal cost of providing more information exceeds the marginal benefit. So, the sender provides the minimum amount of information for the receiver to search, which does not depend on the sender's cost. Hence, the optimal information structure does not depend on the relative cost of information provision.

When the search cost is low, and the relative cost of information provision increases, the marginal cost of providing information increases while the marginal benefit remains the same. Since the sender incurs the information provision cost in the first period for sure, she provides less information in the first period. This allows her to provide more information in the second period when the information provision cost is not always incurred, and she faces information under-provision issues.

## 3.5 A Numerical Example

We illustrate the optimal strategy of the sender by a numerical example. The sender's cost function has the truncated quadratic form, $K(\lambda) = k\lambda^2/(1-\lambda)$, which satisfies Assumption 1. In each figure below, we present the optimal strategy and the sender's payoffs from the optimal one-period, $S_0$, and $S_+$ strategies[7]. The search cost is low in Figure 5. The sender always prefers $S_0$ strategy when she provides information in both periods. As illustrated, the arrival rates of good news at both periods, $\lambda_0^*$ and $\lambda_1^*$, are continuous and increase in $\mu_0$. The beliefs after observing good news at each period, $\bar{\mu}_0^*$ and $\bar{\mu}_1^*$, are continuous and decrease in $\mu_0$. When the prior is lower than the intercept of the brown line, the sender prefers providing information in only one period to smoothening the information provision. When the search cost increases to the intermediate level (Figure 6), the sender always provides information in both periods as long as it is feasible. So, we do not plot the sender's payoff of providing information in only one period. The sender prefers $S_0$ strategy for small prior and switches to $S_+$ strategy as the prior increases. The optimal strategy is non-monotonic and

[7] The domain of the prior is $[c(2v_g - c)/(v_g)^2, \widehat{\mu}_0 \wedge p]$. When $\mu_0 < c(2v_g - c)/(v_g)^2$, the sender provides information in at most one period. When $\mu_0 \geq \widehat{\mu}_0 := 2c - v_b\lambda_1^{**} - [c + (1-\lambda_1^{**})v_b]\lambda_0^{**}$, the sender always chooses the benchmark solution.

discontinuous in $\mu_0$ due to the switch. When the search cost further increases (Figure 7), the optimal information structure is perfectly smooth, and good news always fully reveals the state.

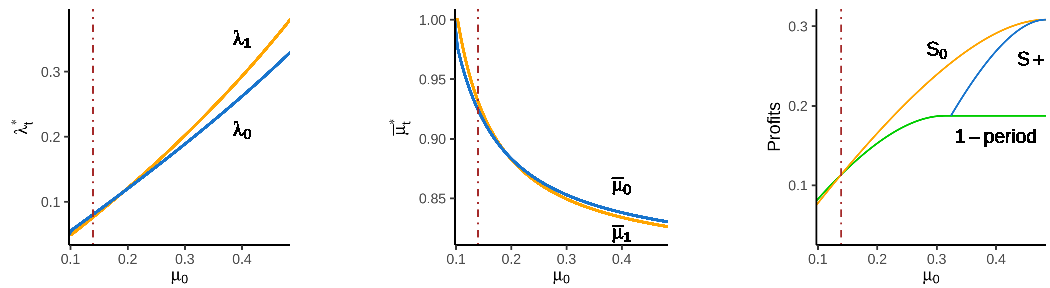


Figure 5: The optimal strategy when $c = 0.01, p = 0.8, v_g = 0.2, v_b = -0.8, k = 0.5$

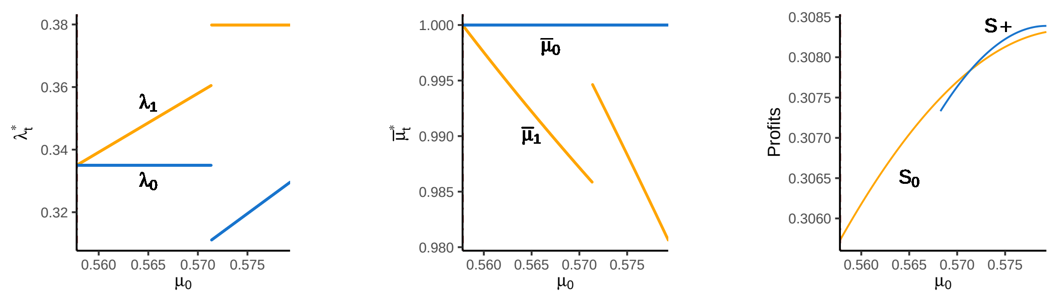


Figure 6: The optimal strategy when $c = 0.067, p = 0.8, v_g = 0.2, v_b = -0.8, k = 0.5$

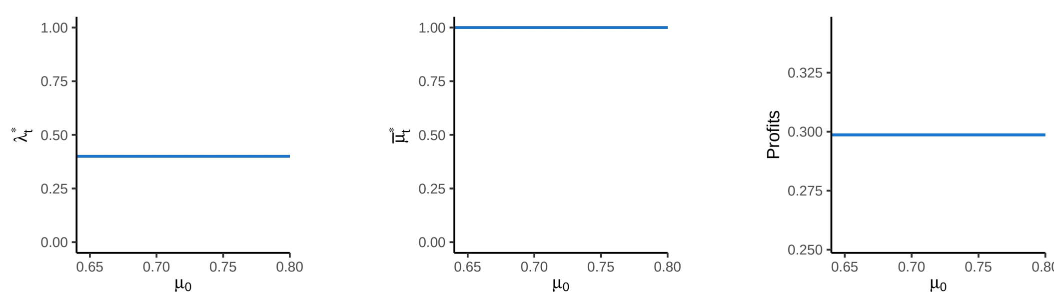


Figure 7: The optimal strategy when $c = 0.08, p = 0.8, v_g = 0.2, v_b = -0.8, k = 0.5$

# 4 The Efficient Information Structure

In all the previous sections, the sender designs the information structure to maximize the expected payoff. This section characterizes the efficient strategy when a social planner designs the information structure to maximize the total welfare. We then investigate the information distortion caused by not taking into account receiver surplus. For tractability reasons, we use a special form of the payoff function, specified in section 2.2.1. So, in this section, $v_g = 1 - p$ and $v_b = -p$.

## 4.1 Efficient Strategy of the Subgame

As discussed in the previous section, the sender does not provide information if the belief at the beginning of the second period is too low, $\mu_1 < \mu_{0,1}$. So, we concentrate on the case in which $\mu_1 \geq \mu_{0,1}$. In the second period, the social planner's problem is:

$$\begin{aligned} &\max_{\lambda_1, \bar{\mu_1}} -K(\lambda_1) + p\lambda_1 + \lambda_1(\bar{\mu_1} - p) - c \qquad (E_1) \\ \text{s.t. } &(IR_1), (F_1) \end{aligned}$$

We discuss below the efficient information structure of the subgame intuitively. The formal characterization of the efficient strategy of the subgame is in the appendix. When the search cost is high, it is very costly for the receiver to search. The sender needs to provide a lot of information to persuade the receiver to search. As a result, the marginal cost of providing more information exceeds the marginal benefit. So, the sender provides the minimum amount of information for the receiver to search, not depending on whether the sender maximizes the sender's surplus or total welfare. Hence, there is no information distortion. When the search cost is lower, it is easier to persuade the receiver to search. The sender provides more than the minimum amount of information for the receiver to search. The marginal costs of providing more information are the same for both the sender and the social planner, while the marginal benefit of providing more information is smaller for the sender. Therefore, the sender provides less information than the social planner does when the search cost is high, or the initial belief is high. One exception is that when both the search cost and the initial belief are low, the sender provides more information than the social planner does because the social planner could only generate low-frequent signals.

## 4.2 Efficient Strategy of the Full Game

Like the previous section, the social planner does not provide information if the prior is too low or the search cost is too high. When she provides information in only one period, the previous subsection (and Proposition 14 in the appendix, where the subscript 1 needs to be replaced by 0) characterized the efficient strategy. When she provides information in both periods, her problem is:

$$\begin{aligned} \max_{\lambda_0,\bar{\mu_0},\mu_1,\lambda_1,\bar{\mu_1}} \ & -K(\lambda_0)+\lambda_0\bar{\mu_0}-c+(1-\lambda_0)\left[-K(\lambda_1)+\lambda_1-c\right] \qquad (E_2) \\ \text{s.t. } & (IR_0),(F_0),(\lambda_1,\bar{\mu}_1) \text{ solves } (E_1) \end{aligned}$$

The following proposition compares the payoff-maximizing strategy and the efficient one when the sender provides information in both periods.

**Proposition 6.** *Suppose the sender provides information in both periods. When $c \geq v_g\lambda_1^{**}$, the payoff-maximizing strategy is efficient. When $c < v_g\lambda_1^{**}$, the sender, who maximizes the payoff, provides less information in the first period and more information in the second period than the social planner does, who maximizes the total welfare.*

When the search cost is high, similar to the argument in the previous subsection, the sender provides the minimum amount of information for the receiver to search, which does not depend on the sender's objective (maximizing the sender surplus or total welfare). Hence, there is no information distortion. When the search cost is lower, it is easier to persuade the receiver to search. The sender provides more than the minimum amount of information for the receiver to search. The marginal costs of providing more information are the same for both the sender and the social planner, while the marginal benefit of providing more information is smaller for the sender. Therefore, the sender provides less information than the social planner does when the search cost is high, or the initial belief is high. One exception is that when both the search cost and the initial belief are low, the sender provides more information than the social planner does because the social planner could only generate low-frequent signals.

# 5 Model Extensions

## 5.1 Dynamic Commitment

Under many circumstances, the assumption that the sender could generate credible signals within each period but does not have dynamic commitment power is reasonable. It is hard for the sender to commit to the entire information structure across all periods ex-ante and convince the receiver that the sender would stick to the information structure when it is profitable to deviate to a different information structure during intermediate periods. However, factors such as reputation could give the sender a stronger commitment power. Here, we consider the case in which the sender has dynamic commitment power and study its implications. She chooses and commits to the entire information structure to maximize the ex-ante expected surplus.

$$\max_{\lambda_0,\bar{\mu_0},\mu_1,\lambda_1,\bar{\mu_1}} -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1) + p\lambda_1\right] \tag{$P_{dc}$}$$
$$\text{s.t. } (IR_0), (F_0), (IR_1), (F_1)$$

**Proposition 7.** *Suppose the search cost is high, $v_g\lambda_1^{**} < c < \widehat{c}$, and $\mu_0 > c(2v_g - c)/(v_g)^2$. The sender provides information in both periods regardless of the dynamic commitment power. If the sender has dynamic commitment power, her payoff is strictly higher, and the receiver gets a strictly positive surplus in the second period. The benefit of dynamic commitment power for the seller vanishes as the search cost approaches zero.*

When the search cost is high, we have shown that the sender would perfectly smoothen the information if she doesn't have dynamic commitment power. If the sender instead has dynamic commitment power, she would commit to providing information more favorable to the receiver in the second period. As a result, the receiver will search even if the sender provides less information in the first period, relaxing the information over-provision issue. Though it hurts the sender's payoff in the second period, it increases the sender's payoff in the first period by reducing the information provision cost. The overall effect is strictly positive. So, the optimal information provision would not be perfectly smooth. This finding is related to the result on durable good pricing. Without dynamic commitment power, the monopolist tends to reduce the price as time goes on, hurting the profit as a rational receiver would strategically wait. Here, dynamic commitment power also benefits the sender, but the underlying mechanisms differ. This paper focuses on persuasion (information provision) rather than incentive (pricing). In addition, in the durable good pricing example, the ability to commit not to provide a more favorable price in the future benefits the sender. While in this paper, the ability to commit to more favorable information in the future benefits the sender.

However, when the search cost approaches zero, the difference between the sender surplus with and without dynamic commitment power approaches zero. The intuition is that the benefit of dynamic commitment power comes from relaxing the participation constraint in the first period by committing to providing more favorable information in the second period. However, the participation constraint is already very loose when the search cost is low. Thus, The benefit of dynamic commitment power approaches zero.

## 5.2 Discounting

In the main model, we assume that there is no discounting. In reality, information acquisition and provision usually happen in a short period, and that assumption is reasonable. However, some communications between the sender and the receiver could take longer. Here, we study the information provision strategy when the sender and the receiver have the same discount factor, $\delta \in (0,1)$. With discounting, the sender's problem becomes:

$$\max_{\lambda_0,\bar{\mu_0},\mu_1,\lambda_1,\bar{\mu_1}} -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\lambda_1) + p\lambda_1\right] \tag{$P_{2,\delta}$}$$

$$\text{s.t. } \lambda_0(\bar{\mu_0} + v_b) + \delta(1-\lambda_0)[\lambda_1(\bar{\mu_1} + v_b) - c] \geq c \tag{$IR_{0,\delta}$}$$

$$(F_0), (\lambda_1, \bar{\mu}_1) \text{ solves } (P_1)$$

One can see that both the objective function and the first-period participation constraint change. When the players become less patient (the discount factor $\delta$ decreases), the present value of the second-period sender surplus decreases, and the first-period participation constraint becomes tighter. Thus, it is less attractive for the sender to sell the goods in the second period.

**Proposition 8.** *When the search cost is high, $c \geq v_g\lambda_1^{**}$, the optimal strategy of the sender does not depend on the discount factor, $\delta$. When the search cost is low, $c \leq \tilde{c}$, the sender provides (weakly) more information in the first period and (weakly) less information in the second period, as the players become less patient.*

When the search cost is high, the participation constraints of the receiver are hard to satisfy. Providing enough information to persuade the receiver to search dominates the force of discounting. Therefore, the sender's strategy remains the same as the no-discounting case, and the sender perfectly smoothens information provision. When the search cost is low, the sender provides (weakly) less information in the second period when the players are less patient because of discounting. The sender provides more information in the first period as she becomes more tempted to convert the receiver early.

### 5.3 Myopic receiver

The assumption that the receiver is forward-looking and takes into account the potential payoff of the second period when he chooses the action of the first period is reasonable if the information environment is transparent and the receiver knows that gradual learning is possible. If the receiver is myopic under some circumstances, then he only trades off the current-period benefit and cost in deciding whether to search or not. The information structure in the second period cannot relax the first-period participation constraint. Therefore, the feasible information structure when the receiver is myopic is a subset of when the receiver is forward-looking. If the receiver surplus in the second period is strictly positive when the receiver is forward-looking, the optimal information structure may not be feasible when the receiver is myopic. Hence, the sender is (weakly) worse off if the receiver is myopic rather than forward-looking. This result has managerial implications as it suggests that the sender should try to better inform the receiver of the possibility of gradual learning. Common knowledge of gradual information revelation improves the sender surplus.

## 6 Concluding Remarks

This paper examines the optimal information provision strategy of a sender and the optimal information acquisition strategy of a receiver when the sender sequentially persuades a receiver to take the desired action. The sender prefers that action regardless of the unknown state, while the receiver only wishes to take that action if the state is good. In our model, the sender incurs a cost to provide information, and the receiver incurs a cost to search. The receiver trades off the cost of searching and the benefit of obtaining more accurate information to make better decisions. The sender trades off the cost of information provision and the benefit of persuading the receiver to search and take the desired action. We allow for sequential information provision and acquisition, so the communication between the receiver and the sender is gradual. Consequently, the sender also makes the intertemporal tradeoff of smoothening the information to reduce the total cost. In equilibrium, she uses one-shot experiments that induce the receiver to take the desired action upon observing good news immediately. The sender smoothens the information provision over multiple periods if and only if the prior on the state being good is high. The sender extracts all the surplus from the receiver when she provides information in both periods while may leave some rent to the receiver when she provides information in only one period. When the search cost of the receiver is high, the receiver is sure that the state is good when he takes the desired action. When the search cost is low, the optimal information structure does not fully reveal the state, which may be bad despite the receiver taking the desired action. We compare the payoff-maximizing information structure with the efficient information structure and find no information distortion when the search cost is high. There could be upward or downward information distortion when the search cost is lower.

We consider a finite-period model in this paper. It is reasonable if there are some kinds of deadlines in the information acquisition phase. If there is no limit on how long the receiver would search, an infinite-period model is more appropriate. In addition, it would be interesting to study the optimal persuasion strategy when there is more than one sender. The competition may lead the sender to provide more information and improve equilibrium efficiency. Moreover, the sender has complete control of the information structure in this paper. It would be interesting to consider the case when the sender could only partially control the information environment.

# References

Aumann, R. J., Maschler, M. (1995). Repeated games with incomplete information. MIT press.

Branco, F., Sun, M., Villas-Boas, J. M. (2012). Optimal search for product information. Management Science, 58(11), 2037-2056.

Branco, F., Sun, M., Villas-Boas, J. M. (2016). Too much information? Information provision and search costs. Marketing Science, 35(4), 605-618.

Che, Y. K., Kim, K., Mierendorff, K. (2020). Keeping the listener engaged: a dynamic model of bayesian persuasion. arXiv preprint arXiv:2003.07338.

Degan, A., Li, M. (2021). Persuasion with costly precision. Economic Theory, 72(3), 869-908.

Ely, J. C. (2017). Beeps. American Economic Review, 107(1), 31-53.

Ely, J. C., Szydlowski, M. (2020). Moving the goalposts. Journal of Political Economy, 128(2), 468-506.

Gentzkow, M., Kamenica, E. (2014). Costly persuasion. American Economic Review, 104(5), 457-62.

Iyer, G., Zhong, Z. Z. (2021). Pushing Notifications as Dynamic Information Design. Marketing Science

Kamenica, E., Gentzkow, M. (2011). Bayesian persuasion. American Economic Review, 101(6), 2590-2615.

Ke, T. T., Villas-Boas, J. M. (2019). Optimal learning before choice. Journal of Economic Theory, 180, 383-437.

Ke, T. T., Shen, Z. J. M., Villas-Boas, J. M. (2016). Search for information on multiple products. Management Science, 62(12), 3576-3603.

Kuksov, D., Villas-Boas, J. M. (2010). When more alternatives lead to less choice. Marketing Science, 29(3), 507-524.

Li, F., Norman, P. (2021). Sequential persuasion. Theoretical Economics

Moscarini, G., Smith, L. (2001). The optimal level of experimentation. Econometrica, 69(6), 1629-1644.

Orlov, D., Skrzypacz, A., Zryumov, P. (2020). Persuading the principal to wait. Journal of Political Economy, 128(7), 2542-2578.

Wei, D. (2021). Persuasion under costly learning. Journal of Mathematical Economics, 94, 102451.

Wu, W. (2021). Sequential bayesian persuasion. Working paper

# Appendix

The following proposition formalizes the claim that we could work with mean-preserving posterior beliefs rather than the specific signal structure in section 2.1.

**Proposition 9.** ***(Equivalent Representations of the Signal)*** *The following are equivalent:*

1. *There exists a binary signal $s \in \Delta(\{0,1\})$ such that $\mathbb{P}[s=1|g] > \mathbb{P}[s=1|b]$.*

2. *There exists a binary-valued posterior belief whose expectation is equal to the prior.*

*Proof of Proposition 9.* $1 \Rightarrow 2$ : Given a binary signal such that $\mathbb{P}[s=1|g] > \mathbb{P}[s=1|b]$, law of iterated expectation implies that $\mathbb{E}[\mathbb{P}[g|s]] = \mathbb{E}[\mathbb{E}[\mathbf{1}_{[g]}|s]] = \mathbb{E}[\mathbf{1}_{[g]}] = \mathbb{P}[g]$. So, the expectation of the posterior belief is equal to the prior. Note that $\mathbb{P}[s=0|g] < \mathbb{P}[s=01|b]$. By Bayes' rule, $\mathbb{P}[g|s=1] = \frac{\mathbb{P}[s=1|g]\mathbb{P}[g]}{\mathbb{P}[s=1|g]\mathbb{P}[g]+\mathbb{P}[s=1|b]\mathbb{P}[b]} > \frac{\mathbb{P}[s=1|g]\mathbb{P}[g]}{\mathbb{P}[s=1|g]\mathbb{P}[g]+\mathbb{P}[s=1|g]\mathbb{P}[b]} = \mathbb{P}[g] = \frac{\mathbb{P}[s=0|g]\mathbb{P}[g]}{\mathbb{P}[s=0|g]\mathbb{P}[g]+\mathbb{P}[s=0|g]\mathbb{P}[b]} > \frac{\mathbb{P}[s=0|g]\mathbb{P}[g]}{\mathbb{P}[s=0|g]\mathbb{P}[g]+\mathbb{P}[s=0|b]\mathbb{P}[b]} = \mathbb{P}[g|s=0]$. So, the posterior belief is binary-valued.

$2 \Rightarrow 1$ : Given a binary-valued posterior belief whose expectation is equal to the prior, $\mu_0$. Denote the distribution of the belief by $\mu = \begin{cases} \bar{\mu}_1 > \mu_0 & w.p.\ \lambda_1 \\ \underline{\mu}_1 < \mu_0 & w.p.\ 1-\lambda_1 \end{cases}$. We now construct a binary signal $s \in \Delta(\{0,1\})$. Define $\mathbb{P}[s=1|g] = \frac{\bar{\mu}_1\lambda_1}{\mu_0}$ and $\mathbb{P}[s=1|b] = \frac{(1-\bar{\mu}_1)\lambda_1}{(1-\mu_0)}$. One can verify by Bayes' rule that this signal $s$ induces exactly the same posterior belief, using the assumption that $\mu_0 = \lambda_1\bar{\mu}_1 + (1-\lambda_1)\underline{\mu}_1$. We just need to show that $\mathbb{P}[s=1|g] > \mathbb{P}[s=1|b]$, which follows from the fact that $\bar{\mu}_1 > \mu_0$. $\square$

*Proof of Proposition 1.* We first characterize the optimal one-period strategy (providing an one-shot experiment) of the sender. Analagous to section 2.3.1, the sender's problem is:

$$\max_{\lambda_0,\bar{\mu_0}} -K(\lambda_0) + p\lambda_0 \tag{$P_0$}$$

$$\text{s.t. } \lambda_0(\bar{\mu_0} + v_b) \geq c \tag{$IR_0'$}$$

$$\lambda_0\bar{\mu_0} + (1-\lambda_0)\underline{\mu}_0 = \mu_0, \tag{$F_0$}$$

$$\lambda_0 \in [0,1], \underline{\mu}_0 \in [0,\mu_0)$$

We transform $(P_0)$ into an equivalent constrained program that is easier to analyze.

**Lemma 2.** *If $\mu_0 < c/v_g$, the sender does not provide information in the second period. If $\mu_0 \geq c/v_g$, $(P_0)$ is equivalent to:*

$$\Pi_1(\mu_0) := \max -K(\lambda_0) + p\lambda_0 \qquad (P_0')$$
$$s.t.\ \lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right]$$

*Proof.* We first show that any $(\lambda_0, \mu_0)$ satisfying the constraints in $(P_0)$, $(IR_0'), (F_0)$, also satisfy the constraints in $(P_0')$:

$(IR_0') \Rightarrow \lambda_0 \geq \frac{c}{\bar{\mu_0}+v_b} \geq \frac{c}{v_g}$. $(IR_0')$ & $(F_0) \Rightarrow \lambda_0 \leq \frac{\mu_0 - c - \underline{\mu}_0}{-v_b - \underline{\mu}_0} \leq \frac{\mu_0 - c}{-v_b}$. Thus, $\lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0-c}{-v_b}\right]$. It is feasible for the sender to provide information in the second period iff $\left[\frac{c}{v_g}, \frac{\mu_0-c}{-v_b}\right]$ is non-empty: $\frac{c}{v_g} \leq \frac{\mu_0-c}{-v_b} \Leftrightarrow \mu_0 \geq \frac{c}{v_g}$. So, If $\mu_0 < \frac{c}{v_g}$, the sender would not provide information in the second period.

We then show that for any $(\lambda_0, \mu_0)$ satisfying the constraints in $(P_0')$ and $\mu_0 \geq \frac{c}{v_g}$, we could find $\bar{\mu_0}, \underline{\mu}_0$ such that $(\lambda_0, \mu_0, \bar{\mu_0}, \underline{\mu}_0)$ satisfies the constraints in $(P_0)$. The conclusion then follows.

Suppose $(\lambda_0, \mu_0)$ satisfies the constraints in $(P_0')$: $\lambda_0 \in \left[\frac{c}{v_g}, \frac{\mu_0-c}{-v_b}\right], \mu_0 \geq \frac{c}{v_g}$. Consider $\bar{\mu_0} = \frac{c}{\lambda_0} - v_b$ and $\underline{\mu}_0 = \frac{\mu_0 - c + v_b\lambda_0}{1-\lambda_0}$. One can verify that $(\lambda_0, \mu_0, \bar{\mu_0}, \underline{\mu}_0)$ satisfies $(IR_0')$ & $(F_0)$. So, we just need to show that $-v_b \leq \bar{\mu_0} \leq 1$ and $\underline{\mu}_0 \geq 0$. $\bar{\mu_0} = \frac{c}{\lambda_0} - v_b \geq -v_b$. $\lambda_0 \geq \frac{c}{v_g} \Rightarrow \bar{\mu_0} = \frac{c}{\lambda_0} - v_b \leq 1$. $\lambda_0 \leq \frac{\mu_0-c}{-v_b} \Rightarrow \mu_0 \geq c - v_b\lambda_0 \Rightarrow \underline{\mu}_0 = \frac{\mu_0 - c + v_b\lambda_0}{1-\lambda_0} \geq 0$. $\square$

Now consider the transformed program $(P_0')$ when $\mu_0 \geq \frac{c}{v_g}$.

1. If $c \geq v_g\lambda_1^{**}$ (*i.e.* $\lambda_1^{**} \leq \frac{c}{v_g}$) and the sender provides information, then $\lambda_0^* = \frac{c}{v_g}$ due to strict concavity of the objective function. One can show that $(\lambda_0, \bar{\mu_0}, \underline{\mu}_0) = (\frac{c}{v_g}, 1, \frac{\mu_0 v_g - c}{v_g - c})$ is the only feasible information structure that satisfies $(IR_0')$ and $(F_0)$. Thus, the sender would provide information with $(\lambda_0, \bar{\mu_0}, \underline{\mu}_0) = (\frac{c}{v_g}, 1, \frac{\mu_0 v_g - c}{v_g - c})$ iff the sender surplus, $-K(\frac{c}{v_g}) + p \cdot \frac{c}{v_g}$, is positive (when it is 0, the sender is indifferent between providing information or not). Let $f(\tilde{c}) = -K(\frac{\tilde{c}}{v_g}) + p \cdot \frac{\tilde{c}}{v_g}$. We have $f(0) = 0$, $f$ is strictly concave and obtains the maximum at $\tilde{c}^* = v_g\lambda_1^{**} < c < 1$. In addition, $f(v_g) < 0$ because $\lim_{\lambda \to 1} K'(\lambda) = +\infty$. Therefore, there exists a unique $\widehat{c} \in (v_g\lambda_1^{**}, v_g)$ $s.t.$ $f(c) \begin{cases} \geq 0, & if\ 0 \leq c \leq \widehat{c} \\ < 0, & if\ c > \widehat{c} \end{cases}$. Moreover, when the sender provides information, $\mu_0 \geq \frac{c}{v_g} \Rightarrow \widehat{c} \leq \mu_0 v_g$. So, the sender does not provide information if $c > \widehat{c}$ and provides information with $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{c}{v_g}, 1)$ if $c < \widehat{c}$. The receiver surplus is 0.

2. If $c \in [\mu_0 + v_b\lambda_1^{**}, v_g\lambda_1^{**})$ (*i.e.* $\lambda_1^{**} \geq \frac{\mu_0-c}{-v_b} > \frac{c}{v_g}$) and the sender provides information, then $\lambda_0^* = \frac{\mu_0-c}{-v_b}$ due to strict concavity of the objective function. One can show

that $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0) = (\frac{\mu_0 - c}{-v_b}, \frac{-\mu_0 v_b}{\mu_0 - c}, 0)$ is the only feasible information structure that satisfies $(IR_0')$ and $(F_0)$. Thus, the sender would provide information with $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0) = (\frac{\mu_0 - c}{-v_b}, \frac{-\mu_0 v_b}{\mu_0 - c}, 0)$ iff the sender surplus, $-K(\frac{\mu_0 - c}{-v_b}) + p \cdot \frac{\mu_0 - c}{-v_b}$, is positive. Since $-K(0) + p \cdot 0 = 0$, $\frac{\mu_0 - c}{-v_b} < \lambda_1^{**}$, and the objective function is strictly concave, the sender surplus is always strictly positive. So, the sender would always provide information. The receiver surplus is 0.

3. If $c < \mu_0 + v_b \lambda_1^{**} \wedge v_g \lambda_1^{**}$ (*i.e.* $\lambda_1^{**} \in \left(\frac{c}{v_g}, \frac{\mu_0 - c}{-v_b}\right)$), then the sender could obtain the maximum possible payoff by setting $(\lambda_0, \bar{\mu}_0) = (\lambda_1^{**}, \frac{\mu_0}{\lambda_1^{**}} \wedge 1)$. Let $\underline{\mu}_0 = \begin{cases} 0, \ if \ \mu_0 \leq \lambda_1^{**} \\ \frac{\mu_0 - \lambda_1^{**}}{1 - \lambda_1^{**}}, \ if \ \mu_0 > \lambda_1^{**} \end{cases}$.
One can verify that $(\lambda_0, \bar{\mu}_0, \underline{\mu}_0)$ is feasible and satisfies $(IR_0')$ and $(F_0)$. We have shown in the proof of Lemma 1 that the sender surplus is strictly positive. So, the sender would provide information and $(\lambda_0^*, \bar{\mu}_0^*) = (\lambda_1^{**}, \frac{\mu_0}{\lambda_1^{**}} \wedge 1)$. The receiver surplus is $\begin{cases} \mu_0 + v_b \lambda_1^{**} - c, \ if \ \mu_0 \leq \lambda_1^{**} \\ \lambda_1^{**} v_g - c, \ if \ \mu_0 > \lambda_1^{**} \end{cases} > 0.$

If $c \geq v_g \lambda_1^{**}$, or $c < v_g \lambda_1^{**}$ and $\mu_1 \leq c - v_b \lambda_1^{**}$, the expected receiver surplus in the second period is 0. The receiver incurs search cost without any immediate benefit in the first period under iterative signals. The expected receiver surplus in the first period is strictly negative if he searches. Therefore, he would not search, and iterative signals are not feasible. Now we consider the case in which $c < v_g \lambda_1^{**}$ and $\mu_1 > c - v_b \lambda_1^{**}$.

The sender's problem when she uses iterative signals is:

$$\Pi_{iter}(\mu_0) := \max -K(\lambda_0) + \lambda_0 \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \tag{$P_{iter}$}$$
$$\text{s.t. } \lambda_1(\bar{\mu}_1 + v_b) \geq \frac{1 + \lambda_0}{\lambda_0} c \tag{$IR_{0,iter}$}$$
$$\lambda_1(\bar{\mu}_1 + v_b) \geq c \tag{$IR_{1,iter}$}$$
$$\lambda_0 \bar{\mu}_0 + (1 - \lambda_0)\underline{\mu}_0 = \mu_0 \tag{$F_0$}$$
$$\lambda_1 \bar{\mu}_1 + (1 - \lambda_1)\underline{\mu}_1 = \mu_1 \tag{$F_1$}$$
$$\mu_1 = \bar{\mu}_0, \lambda_1 = \lambda_1^{**}$$

Note that $(IR_{0,iter})$ implies $(IR_1^{iter})$ and $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}})$ satisfies $(F_1)$. $\forall \mu_1 \geq \lambda_1^{**}$, the optimal second-period strategy is always $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, 1)$. Therefore, choosing $\mu_1$ above $\lambda_1^{**}$ does not increase the second-period sender's payoff or relax the first-period constraints. So, we could restrict $\mu_1$ to be less than or equal to $\lambda_1^{**}$.

i) $\mu_0 \geq c - v_b \lambda_1^{**}$

$\Pi_1(\mu_0) = -K(\lambda_1^{**}) + p\lambda_1^{**} > \Pi_{iter}(\mu_0)$.

ii) $\mu_0 < c - v_b\lambda_1^{**}$

$\Pi_1(\mu_0) = -K(\frac{\mu_0 - c}{-v_b}) + \frac{(\mu_0 - c)p}{-v_b}$

$$\begin{aligned}(IR_{0,iter}) \Rightarrow & \lambda_0 \geq \frac{c}{\mu_1 - c + v_b\lambda_1^{**}} (\Rightarrow \mu_1 \geq \frac{1+\lambda_0}{\lambda_0}c - v_b\lambda_1^{**}) \\ & \geq \frac{c}{v_g\lambda_1^{**} - c} \\ (F_0) \Rightarrow & \lambda_0 = \frac{\mu_0 - \underline{\mu}_0}{\mu_1 - \underline{\mu}_0} \leq \frac{\mu_0}{\mu_1} \overset{(1)}{\leq} \frac{\mu_0}{\frac{1+\lambda_0}{\lambda_0}c - v_b\lambda_1^{**}} \Rightarrow \lambda_0 \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}}\end{aligned} \tag{1}$$

A necessary condition for $\lambda_0$ to be well-defined is:

$$\frac{c}{v_g\lambda_1^{**} - c} \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}} \Leftrightarrow \mu_0 \geq \frac{\lambda_1^{**}c}{v_g\lambda_1^{**} - c} (> \frac{c}{v_g})$$

Therefore, it is feasible for the sender to provide a one-period signal whenever it is feasible to provide iterative signals.

Define $\bar{\Pi}_{iter}(\mu_0) := \max_{0 \leq \lambda_0 \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}}} -K(\lambda_0) + \lambda_0 \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$. One can see that $\bar{\Pi}_{iter}(\mu_0) \geq \Pi_{iter}(\mu_0)$. Let $\lambda_0^*(\mu_0) := \operatorname{argmax}_{0 \leq \lambda_0 \leq \frac{\mu_0 - c}{c - v_b\lambda_1^{**}}} -K(\lambda_0) + \lambda_0 \left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$.

We want to show that:

$$\begin{aligned} & \bar{\Pi}_{iter}(\mu_0) < \Pi_1(\mu_0) \\ \Leftrightarrow & -K(\lambda_0^*(\mu_0)) + \lambda_0^*(\mu_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] < -K(\frac{\mu_0 - c}{-v_b}) + \frac{\mu_0 - c}{-v_b}p \end{aligned} \tag{2}$$

$$\frac{d}{d\lambda_0}\left\{-K(\lambda_0) + \lambda_0\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]\right\} = -K'(\lambda_0) - K(\lambda_1^{**}) + p\lambda_1^{**}$$

$$\Rightarrow \lambda_0^*(\mu_0) = \begin{cases} \frac{\mu_0 - c}{c - v_b\lambda_1^{**}} & if\ \mu_0 \leq \mu_0^t \\ \frac{\mu_0^t - c}{c - v_b\lambda_1^{**}} & if\ \mu_0 > \mu_0^t \end{cases}$$

, where $\lambda_0^t > 0$ is defined by $-K'(\lambda_0^t) - K(\lambda_1^{**}) + p\lambda_1^{**} = 0$, $\mu_0^t = \lambda_0^t(c - v_b\lambda_1^{**}) + c$.

When $\mu_0 \leq \mu_0^t$, $\begin{cases} \lambda_0^*(\mu_0) = \frac{\mu_0 - c}{c - v_b\lambda_1^{**}} > \frac{\mu_0 - c}{-v_b} \Rightarrow -K(\lambda_0^*(\mu_0)) < -K(\frac{\mu_0 - c}{-v_b}) \\ \lambda_0^*(\mu_0)p\lambda_1^{**} = \frac{\mu_0 - c}{c - v_b\lambda_1^{**}}p\lambda_1^{**} < \mu_0 - c \\ -\lambda_0^*(\mu_0)K(\lambda_1^{**}) < 0 \end{cases}$ $\Rightarrow$ (2) holds,

where the first inequality holds because $-v_b > \mu_1 > c - v_b\lambda_1^{**}$.

When $\mu_0 > \mu_0^t$, $\bar{\Pi}_{iter}(\mu_0) = \bar{\Pi}_{iter}(\mu_0^t) < \Pi_1(\mu_0^t) < \Pi_1(\mu_0)$. So, (2) holds.

Thus, we have shown that $\Pi_1(\mu_0) > \bar{\Pi}_{iter}(\mu_0)$ for any $\mu_0$ such that iterative signals are feasible. $\square$

*Proof of Lemma 1.* $\lambda_0^{**}$ and $\lambda_1^{**}$ are determined by the first order conditions: $-K'(\lambda_1^{**}) + p = 0$ and $-K'(\lambda_0^{**}) + p + K(\lambda_1^{**}) - p\lambda_1^{**} = 0$. $-K(0) + p \cdot 0 = 0 \,\&\, -K'(\lambda) + p > 0$ for small $\lambda \Rightarrow -K(\lambda_1^{**}) + p\lambda_1^{**} > 0$. Therefore, $-K'(\lambda_0^{**}) + p + K(\lambda_1^{**}) - p\lambda_1^{**} = 0$ implies that $-K'(\lambda_0^{**}) + p > 0 = -K'(\lambda_1^{**}) + p, \Rightarrow K'(\lambda_0^{**}) < K'(\lambda_1^{**}) \Rightarrow \lambda_0^{**} < \lambda_1^{**}$. $-K(0) + p \cdot 0 + (1-0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] > 0$ and strict concavity (w.r.t. $\lambda_0$) of the objective function imply that $-K(\lambda_0^{**}) + p\lambda_0^{**} + (1-\lambda_0^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] > 0$. $\square$

*Proof of Proposition 2.* It can be proved in the same way as in the proof of Proposition 1 when we derive the optimal one-period strategy of the sender, except that we need to replace the subscript 0 by 1 here. $\square$

*Proof of Proposition 3.*

(1) $v_g\lambda_1^{**} \le c < \widehat{c}$

If the sender provides information in both periods, the constrained program of the sender is:

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] \qquad (P_{2H})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \ge \frac{c}{v_g}$$

We first transform $(P_{2H})$ into an equivalent constrained program that is easier to analyze.

**Lemma 3.** *Suppose $v_g\lambda_1^{**} \le c < \widehat{c}$. If $\mu_{0,1} \le \mu_0 < \frac{2v_g - c}{(v_g)^2}c$, the sender provides information in one period. If $\mu_0 \ge \frac{2v_g - c}{(v_g)^2}c$, $(P_{2H})$ is equivalent to:*

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] \qquad (P'_{2H})$$
$$s.t.\ \lambda_0 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - (1+v_g)c}{-v_b v_g - c}\right]$$

*Proof.* The proof of the equivalence between $(P_{2H})$ and $(P'_{2H})$ is similar to that of Lemma 2.

It is feasible for the sender to provide information at both periods if and only if the domain of $\lambda_1$ is non-empty: $\frac{c}{v_g} \le \frac{v_g\mu_0 - (2-p)c}{pv_g - c} \Leftrightarrow \mu_0 \ge \frac{2v_g - c}{(v_g)^2}c$. $\square$

Denote the optimal $\lambda_0$ without constraints by $\lambda_{0,H}^{**}$. $\lambda_{0,H}^{**} = \underset{\lambda_0}{\operatorname{argmax}} -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right]$. The following lemma summarizes the relative size of $\lambda_{0,H}^{**}$, $\lambda_0^{**}$, and $\frac{c}{v_g}$.

**Lemma 4.** $0 < \lambda_{0,H}^{**} < \lambda_1^{**} \leq \frac{c}{v_g}$.

*Proof.* $\lambda_1^{**} < \frac{c}{v_g}$ is the assumption. F.O.C $\Rightarrow K'(\lambda_{0,H}^{**}) = p + K(\frac{c}{v_g}) - \frac{cp}{v_g}$. From Lemma 1, $K'(\lambda_1^{**}) = p$. $-K(\frac{c}{v_g}) + \frac{cp}{v_g} > 0$ when $c < \widehat{c}$. So, $K'(\lambda_{0,H}^{**}) < K'(\lambda_1^{**}) \Rightarrow \lambda_{0,H}^{**} < \lambda_1^{**}$. $-K'(\lambda_{0,H}^{**}) + p + K(\frac{c}{v_g}) - \frac{cp}{v_g} = 0 \Rightarrow K'(\lambda_{0,H}^{**}) = p + K(\frac{c}{v_g}) - \frac{cp}{v_g} > p - \frac{cp}{v_g} = (1 - \frac{c}{v_g})p > 0$, where the last inequality follows from the assumption that $c < v_g$. Thus, $\lambda_{0,H}^{**} > 0$. $\lambda_1^{**} \leq \frac{c}{v_g}$ is the assumption. $\square$

When it is feasible for the sender to provide information in both periods, $\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c$, Lemma 4 and strict concavity of the objective function imply that the optimal two-period strategy of the sender is $(\lambda_t^*, \bar{\mu_t}^*) = (\frac{c}{v_g}, 1), t = 0, 1$. The sender surplus is $(2 - \frac{c}{v_g})\left[-K(\frac{c}{v_g}) + \frac{cp}{v_g}\right] > -K(\frac{c}{v_g}) + \frac{cp}{v_g}$, the sender surplus of the optimal one-period strategy. Therefore, the sender would always provide information in both periods as long as it is feasible.

(2) $c < v_g\lambda_1^{**}$

If the sender provides information in both periods, we first show that we could restrict the domain of $\mu_1$ to be $\leq \lambda_1^{**}$ given that $c < v_g\lambda_1^{**}$. The intuition is that $\forall \mu_1 \geq \lambda_1^{**}$, the optimal second-period strategy is always $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, 1)$. Therefore, choosing $\mu_1$ above $\lambda_1^{**}$ does not increase the second-period sender's payoff or relax the first-period constraints. So, we could restrict $\mu_1$ to be less than or equal to $\lambda_1^{**}$. Formally, when $\lambda_1^{**} \leq \mu_1 < \mu_0$, the constrained program of the sender is:

$$\begin{aligned}
&\max -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \\
\text{s.t. } &\lambda_0(\bar{\mu_0} + v_b) + (1 - \lambda_0)[\lambda_1^{**}v_g - c] \geq c && (\tilde{IR}_0') \\
&\lambda_0\bar{\mu_0} + (1 - \lambda_0)\mu_1 = \mu_0 && (F_0) \\
&\mu_1 \in [\lambda_1^{**}, \mu_0)
\end{aligned}$$

$(\tilde{IR}_0')$ & $(F_0) \Rightarrow \lambda_0 \leq \frac{-2c + \mu_0 - \mu_1 + v_g\lambda_1^{**}}{v_g\lambda_1^{**} - c - v_b - \mu_1} \leq \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1 - \lambda_1^{**}) - c}$ ("=" when $\mu_1 = \lambda_1^{**}$). $(F_0) \Rightarrow \lambda_0 \geq \frac{\mu_0 - \mu_1}{1 - \mu_1}$. The domain of $\lambda_0$ is non-empty iff $\frac{\mu_0 - \mu_1}{1 - \mu_1} \leq \frac{-2c + \mu_0 - \mu_1 + v_g\lambda_1^{**}}{v_g\lambda_1^{**} - c - v_b - \mu_1} \Leftrightarrow \mu_1 \leq \frac{-2c + v_g\lambda_1^{**} + \mu_0[1 - v_g\lambda_1^{**} + c + v_b]}{v_g - c}$. Therefore, smaller $\mu_1$ means it is more likely for the domain of $\lambda_0$ to be non-empty and larger upper bound of $\lambda_0$. So, the optimal $\mu_1$ would never $\in (\lambda_1^{**}, \mu_0)$. Hence, the optimal strategy of the subgame is $(\lambda_1^*, \bar{\mu}_1^*) = \begin{cases} (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}}) & , \ if \ \mu_1 \in (c - v_b\lambda_1^{**}, \lambda_1^{**}] \\ (\frac{\mu_1 - c}{-v_b}, \frac{-\mu_1 v_b}{\mu_1 - c}) & , \ if \ \mu_1 \in [\frac{c}{v_g}, c - v_b\lambda_1^{**}] \end{cases}$ and the constrained program of the full game is either[8]:

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \tag{$P_{2S_+}$}$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \in [c - v_b\lambda_1^{**}, \lambda_1^{**}]$$

or:

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \tag{$P_{2S_0}$}$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \in [\frac{c}{v_g}, c - v_b\lambda_1^{**}]$$

1. $S_+$ strategy (solution to $(P_{2S_+})$)

**Proposition 10.** *Suppose $c < v_g\lambda_1^{**}$ and $\mu_0 < \widehat{\mu_0} = 2c - v_b\lambda_1^{**} - [c + (1-\lambda_1^{**})v_b]\lambda_0^{**}$. If $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, $(P_{2S_+})$ is feasible with the following solution. $\lambda_0^* = \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$; $\bar{\mu}_0^* = \begin{cases} \frac{(v_b - c)(c - v_b\lambda_1^{**}) - v_b\mu_0}{\mu_0 - 2c + v_b\lambda_1^{**}} \in (-v_b, 1) & , if\ \widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**} \\ 1 & , if\ \widehat{\mu_1}(\mu_0) \geq c - v_b\lambda_1^{**} \end{cases}$; $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}})$; $\mu_1^* = \widehat{\mu_1}(\mu_0) \vee c - v_b\lambda_1^{**}$, where $\widehat{\mu_1}(\mu_0) = \frac{2c - v_b\lambda_1^{**} - (1 + c - v_b\lambda_1^{**} + v_b)\mu_0}{c - v_b - \mu_0}$. The receiver gets zero surplus.*

*Proof.* We first transform $(P_{2S_+})$ into an equivalent constrained program that is easier to analyze.

**Lemma 5.** *Suppose $c < v_g\lambda_1^{**}$ and $\mu_0 \geq \mu_{0,1}$. If $\mu_0 \leq 2c - v_b\lambda_1^{**}$ or $\mu_0 < \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, the sender provides information in one period. If $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, $(P_{2S_+})$ is equivalent to:*

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \tag{$P''_{2S_+}$}$$
$$s.t.\ \lambda_0 \in \left(0, \frac{\mu_0 - 2c + v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c}\right]$$

*Proof.* We first show that, if $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$, $(P_{2S_+})$ is equivalent

[8] We include $\mu_1 = c - v_b\lambda_1^{**}$ in $(P_{2S_+})$ as well to simplify the exposition.

to:

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] \qquad (P'_{2S_+})$$

$$\text{s.t. } \lambda_0 \in \left[\frac{\mu_0-\mu_1}{1-\mu_1}, \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} \wedge \frac{\mu_0-\mu_1}{-v_b-\mu_1}\right]$$

$$\mu_1 \in [c - v_b\lambda_1^{**} \vee \widehat{\mu_1}(\mu_0), \lambda_1^{**}]$$

$$\text{, where } \widehat{\mu_1}(\mu_0) = \frac{2c - v_b\lambda_1^{**} - (1 + c - v_b\lambda_1^{**} + v_b)\mu_0}{c - v_b - \mu_0}$$

$(F_0) \Rightarrow \lambda_0 = \frac{\mu_0-\mu_1}{\bar{\mu_0}-\mu_1} \in \left[\frac{\mu_0-\mu_1}{1-\mu_1}, \frac{\mu_0-\mu_1}{-v_b-\mu_1}\right]$. $(IR_0)$ & $(F_0) \Rightarrow \lambda_0 \leq \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$. For $\lambda_0$ to be positive, we need $\mu_0 > 2c - v_b\lambda_1^{**}$. The domain of $\lambda_0$ is non-empty iff $\frac{\mu_0-\mu_1}{1-\mu_1} \leq \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} \Leftrightarrow \mu_1 \geq \widehat{\mu_1}(\mu_0)$. For $\mu_1 \leq \lambda_1^{**}$, we need $\widehat{\mu_1}(\mu_0) \leq \lambda_1^{**} \Leftrightarrow \mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$. We also have that $\mu_1 = \underline{\mu}_0 < \mu_0$. Thus, the constraints in $(P_{2S_+})$ imply the constraints in $(P'_{2S_+})$.

For any $(\lambda_0, \mu_1)$ satisfying the constraints in $(P'_{2S_+})$, consider $(\lambda_0, \mu_1, \bar{\mu_0} = \frac{\mu_0-(1-\lambda_0)\mu_1}{\lambda_0}, \bar{\mu_1} = \frac{\mu_1}{\lambda_1^{**}} \wedge 1, \underline{\mu}_1 = \frac{\mu_1-\lambda_1^{**}\bar{\mu_1}}{1-\lambda_1^{**}})$. $(IR_0)$ & $(F_0)$ are satisfied by construction. $\bar{\mu_0} = \frac{\mu_0-(1-\lambda_0)\mu_1}{\lambda_0} > \frac{\mu_0-(1-\lambda_0)\mu_0}{\lambda_0} = \mu_0$. $\bar{\mu_0} = \frac{\mu_0-(1-\lambda_0)\mu_1}{\lambda_0} \leq \frac{\mu_0-(1-\frac{\mu_0-\mu_1}{1-\mu_1})\mu_1}{\frac{\mu_0-\mu_1}{1-\mu_1}} = 1$. One can verify that $\bar{\mu_1} \in (-v_b, 1]$, $\underline{\mu}_1 \in [0, -v_b)$. Therefore, the $(\lambda_0, \mu_1, \bar{\mu_0}, \bar{\mu_1}, \underline{\mu}_1)$ we constructed satisfies all the constraints in $(P_{2S_+})$ and is feasible.

Therefore, the two programs are equivalent.

We then show that $(P'_{2S_+})$ is equivalent to $(P''_{2S_+})$. It is clear that the constraints in $(P'_{2S_+})$ imply the constraints in $(P''_{2S_+})$. We now show that for any $\lambda_0 \in \left(0, \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}\right]$, we could find $(\lambda_0, \mu_1)$ that satisfies the constraints in $(P'_{2S_+})$ and is feasible. Since $\lambda_0 = \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$ maximizes the objective function among $\lambda_0 \in \left(0, \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}\right]$ when $\mu_0 < \widehat{\mu_0}$, we only need to verify (by construction) that $\lambda_0 = \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}$ can be obtained.

i) $\widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**}$

   Consider $\mu_1 = c - v_b\lambda_1^{**}, \lambda_0 = \frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, \bar{\mu_0} = \frac{\mu_0-(1-\lambda_0)\mu_1}{\lambda_0} = \frac{(v_b-c)(c-v_b\lambda_1^{**})-v_b\mu_0}{\mu_0-2c+v_b\lambda_1^{**}}$. By construction, $(IR_0)$ and $(F_0)$ are satisfied; $\mu_1$'s constraints are also satisfied. So, we just need to verify that $\bar{\mu_0} \in (p, 1)$. $\bar{\mu_0} < 1 \Leftrightarrow (v_b - c)(c - v_b\lambda_1^{**}) - v_b\mu_0 < \mu_0 - 2c + v_b\lambda_1^{**} \Leftrightarrow \widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**}$, which is the assumption. $\bar{\mu_0} > -v_b \Leftrightarrow c < -v_b(1-\lambda_1^{**})$, which holds because $\mu_0 > 2c - v_b\lambda_1^{**} \Rightarrow c < \frac{1}{2}(\mu_0 + v_b\lambda_1^{**}) \leq \mu_0 + v_b\lambda_1^{**} < -v_b + v_b\lambda_1^{**} = -v_b(1-\lambda_1^{**})$.

ii) $\widehat{\mu_1}(\mu_0) \geq c - v_b\lambda_1^{**}$

Consider $\lambda_0 = \frac{\mu_0 - 2c + v_b \lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c}, \bar{\mu_0} = 1, \mu_1 = \frac{\mu_0 - \lambda_0 \bar{\mu_0}}{1-\lambda_0} = \widehat{\mu_1}(\mu_0)$. By construction, $(IR_0)$ and $(F_0)$ are satisfied; $\mu_1$'s constraints are also satisfied. So, we just need to verify that $\mu_1 = \widehat{\mu_1}(\mu_0) \in [c - v_b\lambda_1^{**}, \lambda_1^{**}]$. $\widehat{\mu_1}(\mu_0) \geq c - v_b \lambda_1^{**}$ is the assumption. $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c} \Rightarrow \widehat{\mu_1}(\mu_0) \leq \lambda_1^{**}$.

□

When $\frac{\mu_0 - 2c + v_b \lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c} \geq \lambda_0^{**} (\Leftrightarrow \mu_0 \geq \widehat{\mu_0})$, the optimal $\lambda_0$ is $\lambda_0^{**}$. When $\frac{\mu_0 - 2c + v_b \lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c} < \lambda_0^{**}$, the optimal $\lambda_0$ is $\frac{\mu_0 - 2c + v_b \lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c}$ due to strict concavity of the objective function (denote it by $J(\lambda_0)$). Since the one-period optimal sender surplus is $-K(\lambda_1^{**}) + p\lambda_1^{**} = J(0)$, $J(\cdot)$ is strictly concave and obtains the unique maximum value at $\lambda_0^{**} > \frac{\mu_0 - 2c + v_b \lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c}$, we have $J(\frac{\mu_0 - 2c + v_b \lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c}) > J(0)$. So, the sender always provides information in both periods when it is feasible ($\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$). We will use this observation in the later proofs. According to the proof of Lemma 5, the receiver always gets zero surpluses when $\mu_0 < \widehat{\mu_0}$. $\mu_1 = \widehat{\mu_1}(\mu_0) \vee c - v_b\lambda_1^{**}$ is the smallest $\mu_1$ that supports $\lambda_0^* = \frac{\mu_0 - 2c + v_b \lambda_1^{**}}{-v_b(1-\lambda_1^{**}) - c}$, which gives the receiver the largest surplus in the first period. So, $\mu_1^* = \widehat{\mu_1}(\mu_0) \vee c - v_b\lambda_1^{**}$. $(F_0) \Rightarrow \bar{\mu}_0^* = \frac{\mu_0 - (1-\lambda_0)\mu_1}{\lambda_0} =$

$$\begin{cases} \frac{(v_b - c)(c - v_b\lambda_1^{**}) - v_b\mu_0}{\mu_0 - 2c + v_b\lambda_1^{**}} \in (p, 1) & , if\ \widehat{\mu_1}(\mu_0) < c - v_b\lambda_1^{**} \\ 1 & , if\ \widehat{\mu_1}(\mu_0) \geq c - v_b\lambda_1^{**} \end{cases}$$

□

2. $S_0$ strategy (solution to $(P_{2S_0})$)

**Proposition 11.** *Suppose $c \leq v_g\lambda_0^{**}$. When $\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c$, $(P'_{2S_0})$ is feasible. $\lambda_0^*$, $\lambda_1^*$, and $\mu_1^*$ are continuous and increasing in $\mu_0$, while $\bar{\mu}_0^*$ and $\bar{\mu}_1^*$ are continuous and decreasing in $\mu_0$, in the solution to $(P'_{2S_0})$. The receiver gets zero surplus at each period.*

*Proof.* We first transform $(P_{2S_0})$ into an equivalent constrained program that is easier to analyze.

**Lemma 6.** *Suppose $c < v_g\lambda_1^{**}$. If $\mu_{0,1} \leq \mu_0 < \frac{2v_g - c}{(v_g)^2}c$, the sender provides information in one period. If $\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c$, $(P_{2S_0})$ is equivalent to:*

$$\max -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \qquad (P'_{2S_0})$$

$$s.t.\ \lambda_0 \in \left[\frac{\mu_0 - \mu_1}{1-\mu_1}, \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}\right]$$

$$\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right]$$

*Proof.* Using the same argument as the proof of Lemma 5, one can show that $(P_{2S_0})$ is equivalent to:

$$\max -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b})+\frac{(\mu_1-c)p}{-v_b}\right] \qquad (P''_{2S_0})$$
$$\text{s.t. } \lambda_0\in\left[\frac{\mu_0-\mu_1}{1-\mu_1},\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}\right]$$
$$\mu_1\in\left[\frac{c}{v_g},\frac{v_g\mu_0-c}{v_g-c}\wedge c-v_b\lambda_1^{**}\right]$$

We just need to show that $(P''_{2S_0})$ is equivalent to $(P'_{2S_0})$.

If $\frac{v_g\mu_0-c}{v_g-c}\le c-v_b\lambda_1^{**}$, $\mu_1$'s constraint becomes $\mu_1\in\left[\frac{c}{v_g},\frac{v_g\mu_0-c}{v_g-c}\right]$. So, $(P''_{2S_0})$ is equivalent to $(P'_{2S_0})$.

If $\frac{v_g\mu_0-c}{v_g-c}> c-v_b\lambda_1^{**}$, denote the solution to $(P'_{2S_0})$ by $(\lambda_0,\mu_1)$.

(a) $(\lambda_0^{**},\lambda_1^{**})$ could be obtained ($\lambda_1=\lambda_1^{**}\Leftrightarrow\mu_1=c-v_b\lambda_1^{**}$) in $(P''_{2S_0})$
$\mu_1\in\left[\frac{c}{v_g},\frac{v_g\mu_0-c}{v_g-c}\wedge c-v_b\lambda_1^{**}\right]$ is equivalent to $\mu_1\in\left[\frac{c}{v_g},\frac{v_g\mu_0-c}{v_g-c}\right]$, as the optimal $\mu_1$ under the latter (relaxed) constraint would be $c-v_b\lambda_1^{**}$. So, $(P''_{2S_0})$ is equivalent to $(P'_{2S_0})$.

(b) $(\lambda_0^{**},\lambda_1^{**})$ could not be obtained in $(P''_{2S_0})$
Suppose $\mu_1>c-v_b\lambda_1^{**}$. If $\lambda_0>\frac{\mu_0-\mu_1}{1-\mu_1}$, consider $(\lambda_0'=\lambda_0,\mu_1'=\mu_1-\varepsilon)$. For small enough $\varepsilon$, it is feasible and gives the sender a strictly higher payoff. A contradiction! If $\lambda_0=\frac{\mu_0-\mu_1}{1-\mu_1}$ instead, we have $\frac{\mu_0-\mu_1}{1-\mu_1}\ge\lambda_0^{**}$. A contradiction!
Therefore, $\mu_1\le c-v_b\lambda_1^{**}$ and thus $(P''_{2S_0})$ is equivalent to $(P'_{2S_0})$.

In sum, $(P''_{2S_0})$ is equivalent to $(P'_{2S_0})$. □

To solve $(P'_{2S_0})$, we introduce another lemma:

**Lemma 7.** *Suppose $c<v_g\lambda_1^{**}$. For $\mu_0<\widehat{\mu_0}$, $\lambda_0$ is binding at the upper bound in the solution to $(P'_{2S_0})$.*

*Proof.* To solve $(P'_{2S_0})$, we consider several cases.

i) $\lambda_0\le\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}$ is binding and $\mu_1$'s constraints are not binding.
The Lagrangian is $\mathcal{L}=-K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b})+\frac{(\mu_1-c)p}{-v_b}\right]+\eta\left(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}-\lambda_0\right)$
*s.t.* $\eta\ge0,\eta\left(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}-\lambda_0\right)=0$.

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{v_b} - \eta = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1 - c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \eta \cdot \frac{\mu_0 + v_b - c}{(v_b + \mu_1)^2} = 0 \end{cases}$$

Plug in $\lambda_0 = \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}$. Dividing the second equality by $\frac{\mu_0 + v_b - c}{(v_b + \mu_1)^2}$ and comparing with the first equality, we obtain:

$$\begin{aligned} \eta &= -(v_b + \mu_1)\left[K'(\frac{\mu_1 - c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] \\ &= -K'(\lambda_0) + p + K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{v_b} \\ \Rightarrow \; & K(\frac{\mu_1 - c}{-v_b}) + \frac{v_b + \mu_1}{v_b} K'(\frac{\mu_1 - c}{-v_b}) - K'(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) - \frac{cp}{v_b} = 0 \end{aligned} \tag{$*$}$$

$\frac{\partial}{\partial \mu_1}\left[K(\frac{\mu_1 - c}{-v_b}) + \frac{v_b + \mu_1}{v_b} K'(\frac{\mu_1 - c}{-v_b})\right] = -\frac{v_b + \mu_1}{(v_b)^2} K''(\frac{\mu_1 - c}{-v_b}) > 0$. So, the sum of the first two terms of the LHS of $(*)$ is strictly increasing in $\mu_1$. $\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}$ is strictly decreasing in $\mu_1$, $K'(\cdot)$ is strictly increasing in $\mu_1$. So, $-K'(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1})$ is strictly increasing in $\mu_1$. Thus, the LHS of $(*)$ is strictly increasing in $\mu_1$. When $\mu_0$ increases, the LHS of $(*)$ is strictly negative if $\mu_1$ is unchanged. Therefore, $\mu_1$ also has to increase. So, the sum of the first two terms of the LHS of $(*)$ increases. As a result, the third term, $-K'(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) = -K'(\lambda_0)$ has to decrease strictly. So, $\lambda_0$ has to increase strictly. In sum, the optimal $\lambda_0$ and $\mu_1$ are strictly increasing in $\mu_0$.

ii) $\mu_1 \leq \frac{v_g \mu_0 - c}{v_g - c}$ is binding.

When $\mu_1 = \frac{v_g \mu_0 - c}{v_g - c}$, $\lambda_0 \in \{\frac{c}{v_g}\}$. So, $\lambda_0$ is binding at the upper bound.

iii) $\mu_1 \geq \frac{c}{v_g}$ is binding and $\lambda_0$ is not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] + \eta\left(\mu_1 - \frac{c}{v_g}\right)$ $s.t.\ \eta \geq 0, \eta\left(\mu_1 - \frac{c}{v_g}\right) = 0$.

$$\text{F.O.C.} \Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{v_b} = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1 - c}{-v_b}) \cdot \frac{1}{v_b} - \frac{p}{v_b}\right] + \eta = 0 \end{cases}$$

The second equality $\Rightarrow \eta = -\frac{1-\lambda_0}{v_b}\left[K'(\frac{\mu_1 - c}{-v_b}) - p\right] \overset{c < v_g \lambda_1^{**}}{<} -\frac{1-\lambda_0}{v_b}\left[K'(\lambda_1^{**}) - p\right] = 0$. But $\eta \geq 0$. A contradiction! So, this case cannot happen.

iv) $\mu_1 \geq \frac{c}{v_g}$ is binding and $\lambda_0 \geq \frac{\mu_0 - \mu_1}{1 - \mu_1}$ is binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] + \eta\left(\mu_1 - \frac{c}{v_g}\right) + \xi\left(\lambda_0 - \frac{\mu_0 - \mu_1}{1 - \mu_1}\right)$ $s.t.\ \eta \geq 0, \eta\left(\mu_1 - \frac{c}{v_g}\right) = 0, \xi \geq 0, \xi\left(\lambda_0 - \frac{\mu_0 - \mu_1}{1 - \mu_1}\right) = 0$.

F.O.C. $\Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{v_b} + \xi = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b})\cdot\frac{1}{v_b} - \frac{p}{v_b}\right] + \eta + \xi\frac{1-\mu_0}{(1-\mu_1)^2} = 0 \end{cases}$

Similar to the previous case, the LHS of the second equality $> 0$. A contradiction! So, this case cannot happen.

v) $\lambda_0 \geq \frac{\mu_0-\mu_1}{1-\mu_1}$ is binding and $\mu_1$ is not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0)+p\lambda_0+(1-\lambda_0)\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right]+\xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right)$

$s.t.\ \xi \geq 0, \xi\left(\lambda_0 - \frac{\mu_0-\mu_1}{1-\mu_1}\right) = 0.$

F.O.C. $\Rightarrow \begin{cases} -K'(\lambda_0) + p + K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b} + \xi = 0 \\ (1-\lambda_0)\left[K'(\frac{\mu_1-c}{-v_b})\cdot\frac{1}{v_b} - \frac{p}{v_b}\right] + \xi\frac{1-\mu_0}{(1-\mu_1)^2} = 0 \end{cases}$

Similar to the previous case, the LHS of the second equality $> 0$. A contradiction! So, this case cannot happen.

vi) both $\lambda_0$ and $\mu_1$ are not binding.

The solution is the unconstrained optimal solution $(\lambda_0^{**}, \lambda_1^{**})$. But we have assumed that it is not feasible.

i) to vi) finish the proof of Lemma 7. $\square$

According to Lemma 7, if $\mu_0 \geq \frac{2v_g-c}{(v_g)^2}c$, $(P'_{2S_0})$ is equivalent to:

$$\max -K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p\,\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} + (1-\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] \tag{$P'''_{2S_0}$}$$

s.t. $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c}\right]$

The first order derivative of the objective function w.r.t. $\mu_1$ is:

$$\begin{aligned} &D(\mu_0,\mu_1) \\ :=&\frac{\partial}{\partial\mu_1}\left\{-K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p\,\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} + (1-\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right]\right\} \\ =&\frac{\mu_0+v_b-c}{(\mu_1+v_b)^2}\left[K(\frac{\mu_1-c}{-v_b}) + \frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b}) - K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) - \frac{cp}{v_b}\right] \end{aligned}$$

The first term of $D(\mu_0,\mu_1)$, $\frac{\mu_0+v_b-c}{(\mu_1+v_b)^2}$, is always strictly negative. The second term, $K(\frac{\mu_1-c}{-v_b}) + \frac{v_b+\mu_1}{v_b}K'(\frac{\mu_1-c}{-v_b}) - K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) - \frac{cp}{v_b}$, is the LHS of $(*)$, which has been shown to be strictly increasing in $\mu_1$ in the proof of Lemma 7. One can see that $D(\mu_0, \mu_1)$ is strictly negative when $\mu_1$ is large. Thus, $D(\mu_0, \mu_1)$ is always negative or positive for $\mu_1$ small and negative for $\mu_1$ large. Let $\mu_1^{**}(\mu_0)$ be the cutoff value such that $D(\mu_0, \mu_1) \geq 0$ for $\mu_1 \leq \mu_1^{**}(\mu_0)$ and $D(\mu_0, \mu_1) \leq 0$ for $\mu_1 \geq \mu_1^{**}(\mu_0)$ ($\mu_1^{**}(\mu_0) := -\infty$ if $D(\mu_0, \mu_1)$ is always negative). Since $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right]$, the optimal $\mu_1^*(\mu_0) = \left[\frac{c}{v_g} \vee \mu_1^{**}(\mu_0)\right] \wedge \frac{v_g\mu_0 - c}{v_g - c}$.

One can see that we could define $\tilde{\mu}_1^{**}(\mu_0) := \begin{cases} \frac{c}{v_g}, \ if\ D(\mu_0, \mu_1) \text{ is always negative} \\ \mu_1^{**}(\mu_0), \text{ otherwise} \end{cases}$.

$\tilde{\mu}_1^{**}(\mu_0) \in (-\infty, +\infty)$ and $\mu_1^*(\mu_0) = \left[\frac{c}{v_g} \vee \tilde{\mu}_1^{**}(\mu_0)\right] \wedge \frac{v_g\mu_0 - c}{v_g - c}$. Since $\tilde{\mu}_1^{**}(\mu_0)$ is continuous in $\mu_0$, $\mu_1^*(\mu_0)$ is also continuous in $\mu_0$. It then implies that $\lambda_0^*(\mu_0) = \frac{\mu_0 - \mu_1^*(\mu_0) - c}{-v_b - \mu_1^*(\mu_0)}$, $\lambda_1^*(\mu_0) = \frac{\mu_1^*(\mu_0) - c}{-v_b}$, and $\bar{\mu}_1^*(\mu_0) = \frac{-\mu_1^*(\mu_0)v_b}{\mu_1^*(\mu_0) - c}$ are continuous in $\mu_0$.

We have shown in the proof of Lemma 7 that $\lambda_0^*$ and $\mu_1^*$ are strictly increasing in $\mu_0$ when $\mu_1^*$ is the interior solution. Now we consider the case when $\mu_1$ is binding. When $\mu_1^* = \frac{c}{v_g}$, $\lambda_0^* = \frac{\mu_0 - \mu_1^* - c}{-v_b - \mu_1^*} = \frac{\mu_0 - \frac{c}{v_g} - c}{-v_b - \frac{c}{v_g}}$ is strictly increasing in $\mu_0$. When $\mu_1^* = \frac{v_g\mu_0 - c}{v_g - c}$, it is strictly increasing in $\mu_0$ and $\lambda_0^* = \frac{\mu_0 - \mu_1^* - c}{-v_b - \mu_1^*} = \frac{c}{v_g}$ is constant. Together with the continuity property we just established, we have shown that $\lambda_0^*$ and $\mu_1^*$ are (weakly) increasing in $\mu_0$. Thus, $\lambda_1^* = \frac{\mu_1^* - c}{-v_b}$ is (weakly) increasing in $\mu_0$ and $\bar{\mu}_1^* = \frac{-\mu_1^* v_b}{\mu_1^* - c}$ is (weakly) decreasing in $\mu_0$. □

According to the proof of Proposition 1 and Lemma 6, the sender does not provide information iff $\mu_0 < \mu_{0,1}$ and provide information in one period if $\mu_{0,1} \leq \mu_0 < \frac{2v_g - c}{(v_g)^2}c$. Thus, we just need to determine whether she provides information in one period or in both periods when $\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c$ by comparing the sender surplus of the optimal one-period strategy and the optimal sender surplus of the $S_0$ strategy.

(a) $c \leq v_g\lambda_0^{**}$

Define $\mu_{1,2} := \inf\{\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c : \Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)\}$. One can see that $\mu_{1,2} \in [\frac{2v_g - c}{(v_g)^2}c, \widehat{\mu_0})$ and $\Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2})$. According to Lemma 7,

$$\Pi_1(\mu_0) = -K(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**}) + p(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**})$$

$$\begin{aligned} &\Pi_{S_0}(\mu_0) \\ =& \max_{\mu_1} -K(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) + p\ \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} + (1 - \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}) \left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] \\ &s.t\ \mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right] \end{aligned}$$

i) $\mu_{1,2} \geq c - v_b\lambda_1^{**}$

$\forall \mu_0 \in (\mu_{1,2}, \widehat{\mu_0}], \Pi_{S_0}(\mu_0) > \Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2}) = \Pi_1(\mu_0)$.

ii) $\mu_{1,2} < c - v_b\lambda_1^{**}$

$\forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**})$, $\frac{d\Pi_1(\mu_0)}{d\mu_0} = K'(\frac{\mu_0 - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b}$.

A. $\mu_1(\mu_0) = \mu_1^u(\mu_0) = \frac{v_g\mu_0 - c}{v_g - c}$

$\Pi_{S_0}(\mu_0) = -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})[-K(\frac{\mu_1^u(\mu_0)-c}{-v_b}) + \frac{(\mu_1^u(\mu_0)-c)p}{-v_b}]$.

For $\Delta > 0$ small enough, we have $\mu_0 + \delta < c - v_b\lambda_1^{**}$, $\forall \delta \in (0, \Delta)$. Consider $\mu_{0,\delta} = \mu_0 + \delta \in (\mu_0, \mu_0 + \Delta)$, we have $\Pi_{S_0}(\mu_{0,\delta}) \geq \underline{\Pi_{S_0}}(\mu_{0,\delta}) := -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})[-K(\frac{\mu_1^u(\mu_{0,\delta})-c}{-v_b}) + \frac{(\mu_1^u(\mu_{0,\delta})-c)p}{-v_b}]$. Noticing that $\Pi_{S_0}(\mu_0) = \underline{\Pi_{S_0}}(\mu_0)$, we have

$$\frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi_{S_0}}(\mu_0)}{d\mu_0} = K'(\frac{\mu_1^u(\mu_0) - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b} > \frac{d\Pi_1(\mu_0)}{d\mu_0}$$

So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0), \forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**})$, and the inequality is strict when $\mu_0 > \mu_{1,2}$.

B. $\mu_1(\mu_0) < \mu_1^u(\mu_0)$

Let $\lambda_0 = \frac{\mu_0 - \mu_1(\mu_0) - c}{-v_b - \mu_1(\mu_0)}$. For $\Delta > 0$ small enough, we have $\mu_0 + \delta < c - v_b\lambda_1^{**}$ and $\mu_1(\mu_0) + \frac{\delta}{1-\lambda_0} < \mu_1^u(\mu_0) < \mu_1^u(\mu_0 + \delta)$, $\forall \delta \in (0, \Delta)$. Consider $\mu_{0,\delta} = \mu_0 + \delta \in (\mu_0, \mu_0 + \Delta)$, noticing that $\frac{\mu_0 + \delta - (\mu_1 + \frac{\delta}{1-\lambda_0}) - c}{-v_b - (\mu_1 + \frac{\delta}{1-\lambda_0})} = \lambda_0$, we have $\Pi_{S_0}(\mu_{0,\delta}) \geq \widetilde{\underline{\Pi_{S_0}}}(\mu_{0,\delta}) := -K(\lambda_0) + p\lambda_0 + (1 - \lambda_0)[-K(\frac{\mu_1(\mu_0) + \frac{\delta}{1-\lambda_0} - c}{p}) + \frac{(\mu_1(\mu_0) + \frac{\delta}{1-\lambda_0} - c)p}{-v_b}]$.

$$\begin{aligned}
\frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} \geq& \frac{d\widetilde{\underline{\Pi_{S_0}}}(\mu_0)}{d\mu_0} \\
=&(1 - \lambda_0)\left[-\frac{1}{p}K'(\frac{\mu_1(\mu_0) - c}{p})\frac{1}{1 - \lambda_0} + \frac{1}{1 - \lambda_0}\right] \\
=&1 - \frac{1}{p}K'(\frac{\mu_1(\mu_0) - c}{p}) \\
\geq&1 - \frac{1}{p}K'(\frac{\mu_1^u(\mu_0) - c}{p}) \\
>&\frac{d\Pi_1(\mu_0)}{d\mu_0}
\end{aligned}$$

So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)$ and the inequality is strict when $\mu_0 > \mu_{1,2}$.

In sum, $\Pi_{S_0}(\mu_0) > \Pi_1(\mu_0)$, $\forall \mu_0 \in (\mu_{1,2}, \widehat{\mu_0}]$.

(b) $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$

The following lemma provides a closed-form solution to program $(P_{2S_0})$ when the search cost is intermediate:

**Lemma 8.** *Suppose* $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$. $\lambda_0^* = \frac{c}{v_g}, \mu_1^* = \frac{v_g\mu_0 - c}{v_g - c}$ *in the solution to* $(P_{2S_0})$.

*Proof.* Fixing a $\mu_1$, consider the unconstrained program:

$$\lambda_0' \in \operatorname{argmax} -K(\lambda_0) + p\lambda_0 + (1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right]$$

Since $-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b} \leq -K(\lambda_1^{**}) + p\lambda_1^{**}$, $\lambda_0' \leq \lambda_0^{**} < \frac{c}{v_g}$. Since $\lambda_0^* = \frac{\mu_0-\mu_1-c}{-v_b-\mu_1}$ is decreasing in $\mu_1$ and $-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}$ is increasing in $\mu_1$ for $\mu_1 < c - v_b\lambda_1^{**}$, it is always better for the sender to choose the largest possible $\mu_1$ ($\lambda_0 = \frac{c}{v_g}$ when $\mu_1$ is the upper bound) when it is less than $c - v_b\lambda_1^{**}$, which is equivalent to $\frac{v_g\mu_0-c}{v_g-c} < c - v_b\lambda_1^{**} \Leftrightarrow \mu_0 < \widehat{\widehat{\mu}}_0 := \frac{c+(v_g-c)(c-v_b\lambda_1^{**})}{v_g}$. $\widehat{\mu}_0 = 2c - v_b\lambda_1^{**} + p\lambda_0^{**}(p - c - p\lambda_1^{**}) < 2c - v_b\lambda_1^{**} + p\cdot\frac{c}{v_g}(p - c - p\lambda_1^{**}) = \widehat{\widehat{\mu}}_0$. Thus, one can see that for $\mu_0 \leq \widehat{\mu}_0$, $\lambda_0^* = \frac{c}{v_g}, \mu_1^* = \frac{v_g\mu_0-c}{v_g-c}$ in the solution to $(P_{2S_0}')$. Lemma 6 shows that $(P_{2S_0}')$ is equivalent to $(P_{2S_0})$. So, $\lambda_0^* = \frac{c}{v_g}, \mu_1^* = \frac{v_g\mu_0-c}{v_g-c}$ are also the solutions to $(P_{2S_0})$. □

Define $\mu_{1,2} := \inf\{\mu_0 \geq \frac{2v_g-c}{(1-p)^2}c : \Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0) \text{ or } S_+ \text{ strategy is feasible}\}$. Note that $\mu_{1,2} \leq \mu_{2,+}$. If $S_+$ strategy is feasible $\forall\mu_0 > \mu_{1,2}$, the 1-period sender surplus would always be dominated by the 2-period sender surplus $\forall\mu_0 > \mu_{1,2}$, as the optimal $S_+$ strategy generates a strictly higher sender surplus than the optimal 1-period strategy. We now consider the case in which $S_+$ strategy is not feasible for some $\mu_0 > \mu_{1,2}$, which implies that $\Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2})$.

$$\begin{aligned}
\Pi_1(\mu_0) =& -K(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**}) + p(\frac{\mu_0 - c}{-v_b} \wedge \lambda_1^{**}) \\
\Pi_{S_0}(\mu_0) =& -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1 - \frac{c}{v_g})\left[-K(\frac{\mu_1^u(\mu_0) - c}{-v_b}) + \frac{(\mu_1^u(\mu_0) - c)p}{-v_b}\right]
\end{aligned}$$

i) $\mu_{1,2} \geq c - v_b\lambda_1^{**}$
   $\forall\mu_0 \in (\mu_{1,2}, \mu_{2,+}], \Pi_{S_0}(\mu_0) > \Pi_{S_0}(\mu_{1,2}) \geq \Pi_1(\mu_{1,2}) = \Pi_1(\mu_0)$.

ii) $\mu_{1,2} < c - v_b\lambda_1^{**}$

$\forall \mu_0 \in [\mu_{1,2}, c - v_b\lambda_1^{**}]$,

$$\frac{d\Pi_1(\mu_0)}{d\mu_0} = K'(\frac{\mu_0 - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b}$$
$$\frac{d\Pi_{S_0}(\mu_0)}{d\mu_0} = K'(\frac{\mu_1^u(\mu_0) - c}{-v_b})\frac{1}{v_b} - \frac{p}{v_b} > \frac{d\Pi_1(\mu_0)}{d\mu_0}$$

So, $\Pi_{S_0}(\mu_0) \geq \Pi_1(\mu_0)$ and the inequality is strict when $\mu_0 > \mu_{1,2}$. $\Pi_{S_0}(c - v_b\lambda_1^{**}) > \Pi_1(c - v_b\lambda_1^{**})$. $\forall \mu_0 \in [c - v_b\lambda_1^{**}, \mu_{2,+}], \Pi_{S_0}(\mu_0) \geq \Pi_{S_0}(c - v_b\lambda_1^{**}) > \Pi_1(c - v_b\lambda_1^{**}) = \Pi_1(\mu_0)$.

In sum, $\forall \mu_0 \in (\mu_{1,2}, \mu_{2,+}], \Pi_{S_0}(\mu_0) > \Pi_1(\mu_0)$.

One can see that the optimal $S_+$ strategy always generates a strictly higher (and strictly positive) sender surplus than the optimal 1-period strategy. Therefore, the sender always provides information in both periods when $S_+$ strategy is feasible.[9] By Lemma 5, $S_+$ strategy is feasible iff $\mu_0 > 2c - v_b\lambda_1^{**}$ and $\mu_0 \geq \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}$ when $c < v_g\lambda_1^{**}$. Hence, together with the above results on $S_0$ strategy, there exists $\mu_{1,2} \in [\frac{2v_g - c}{(v_g)^2}c, 2c - v_b\lambda_1^{**} \vee \frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c}]$ such that the sender does not provide information if $\mu_0 < \mu_{0,1}$, provides information in one period if $\mu_0 \in [\mu_{0,1}, \mu_{1,2})$, and provides information in both periods if $\mu_0 > \mu_{1,2}$. □

*Proof of Proposition 4.*

(1) High Search Cost ($v_g\lambda_1^{**} \leq c < \widehat{c}$)

It has been shown in the proof of Proposition 3. Since the optimal strategy does not depend on the prior, the sender's payoff does not depend on the prior either.

(2) Low Search Cost ($c \leq \tilde{c} = v_g K'^{-1}\left[\frac{K(\lambda_1^{**})}{\lambda_1^{**}}\right]$)

We first prove the following claim:

**Lemma 9.** $\tilde{c} < v_g\lambda_0^{**} < v_g\lambda_1^{**}$.

[9] But the optimal 2-period strategy may be either $S_+$ or $S_0$ strategy.

*Proof.* We have $K'(\frac{\tilde{c}}{v_g}) = \frac{K(\lambda_1^{**})}{\lambda_1^{**}}$.

$$\begin{aligned}
\text{FOC of } (P_b) \Rightarrow & K'(\lambda_0^{**}) = K(\lambda_1^{**}) + p(1-\lambda_1^{**}) \\
\Rightarrow & \lambda_1^{**} K'(\lambda_0^{**}) - K(\lambda_1^{**}) = (1-\lambda_1^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] > 0 \\
\Rightarrow & K'(\lambda_0^{**}) > \frac{K(\lambda_1^{**})}{\lambda_1^{**}} = K'(\frac{\tilde{c}}{v_g}) \\
\Rightarrow & \lambda_0^{**} > \frac{\tilde{c}}{v_g} \\
\Rightarrow & \tilde{c} < v_g \lambda_0^{**} < v_g \lambda_1^{**}
\end{aligned}$$

□

We then compare the optimal sender surplus between the solution to $(P_{2S_+})$ and the solution to $(P_{2S_0})$, and show that the optimal $S_0$ strategy is always preferred to the optimal $S_+$ strategy when both types of strategy are feasible.

**Proposition 12.** *Suppose $c \leq \tilde{c}$ and $\mu_0 < \widehat{\mu_0}$. The sender uses $S_0$ strategy when she provides information in both periods.*

*Proof.* $\forall \mu_0 < \widehat{\mu_0}$ such that $S_0$ ($S_+$) strategy is feasible, denote the optimal sender surplus by $\Pi_{S_0}(\mu_0)$ ($\Pi_{S_+}(\mu_0)$).

i) $\mu_0' := \frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \leq \mu_0 < \widehat{\mu_0}$

   $\frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \leq \mu_0 \Leftrightarrow \widehat{\mu_1}(\mu_0) \leq c - v_b\lambda_1^{**}$. According to Proposition 10, $(\lambda_{0,S_+}, \mu_{1,S_+}) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, c - v_b\lambda_1^{**})$ gives $\Pi_{S_+}(\mu_0)$. $\mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c} \wedge c - v_b\lambda_1^{**}\right]$ in $(P''_{2S_0})$ and $\frac{(v_g-c)(c-v_b\lambda_1^{**})+c}{v_g} \leq \mu_0 \Leftrightarrow \frac{v_g\mu_0-c}{v_g-c} \geq c - v_b\lambda_1^{**}$. Consider $(\lambda_0, \mu_1) = (\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}, c - v_b\lambda_1^{**})$, which satisfies the costraints in $(P''_{2S_0})$ and is identical to $(\lambda_{0,S_+}, \mu_{1,S_+})$. So, $\Pi_{S_0}(\mu_0) \geq \Pi_{S_+}(\mu_0)$.

ii) $\mu_0 < \mu_0'$

   $\mu_0 < \mu_0' \Leftrightarrow \widehat{\mu_1}(\mu_0) > c - v_b\lambda_1^{**}$. According to Proposition 10, $(\lambda_{0,S_+}, \mu_{1,S_+}) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, \widehat{\mu_1}(\mu_0))$ gives $\Pi_{S_+}(\mu_0)$. One can verify that $\frac{v_g\mu_0'-c}{v_g-c} = c - v_b\lambda_1^{**}$, $\frac{\mu_0'-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c} = \frac{c}{v_g} \leq \lambda_0^{**}$. So, $\mu_0' \leq \widehat{\mu_0}$, which implies that $\lambda_{0,S_+}(\mu_0') \leq \lambda_0^{**}$.

   First, consider $\mu_0$ such that both $S_0$ and $S_+$ strategies are feasible. Let $\mu_1^u(\mu_0) := \frac{v_g\mu_0-c}{v_g-c}$.

$$\Pi_{S_+}(\mu_0) = -K(\lambda_{0,S_+}(\mu_0)) + p\lambda_{0,S_+}(\mu_0) + (1-\lambda_{0,S_+}(\mu_0))[-K(\lambda_1^{**}) + p\lambda_1^{**}]$$

$$\Pi_{S_0}(\mu_0) \geq \underline{\Pi}_{S_0}(\mu_0) := -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1-\frac{c}{v_g})\left[-K(\frac{\mu_1^u(\mu_0)-c}{-v_b}) + \frac{(\mu_1^u(\mu_0)-c)p}{-v_b}\right]$$

**Lemma 10.** $\frac{\mu_1^u(\mu_0)-c}{-v_b} \geq \lambda_{0,S_+}(\mu_0)$, $\forall \mu_0 < \mu_0'$.

*Proof.* $\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] = \frac{v_g}{-v_b(v_g-c)}$, $\frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)] = \frac{1}{-v_b(1-\lambda_1^{**})-c}$.

$$\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] \leq \frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)] \Leftrightarrow c(-2v_b-1) \leq -v_b v_g \lambda_1^{**} \qquad (*)$$

If $v_b \geq -1/2$, $(*)$ always holds. If $(-1 <)\ v_b < -1/2$, we have that $\frac{-v_b}{-2v_b-1} \geq 1 \Rightarrow c \leq v_g\lambda_1^{**} \leq \frac{-v_b v_g}{-2v_b-1}\lambda_1^{**} \Rightarrow (*)$ also holds. So, $\frac{d}{d\mu_0}[\frac{\mu_1^u(\mu_0)-c}{-v_b}] \leq \frac{d}{d\mu_0}[\lambda_{0,S_+}(\mu_0)]$, $\forall \mu_0 < \mu_0'$.

Note that $\frac{\mu_1^u(\mu_0')-c}{-v_b} = \lambda_1^{**} > \lambda_0^{**} \geq \lambda_{0,S_+}(\mu_0')$. This concludes the proof. $\square$

Now we calculate the increasing rate of the sender surplus as a function of $\mu_0$:

$$\frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} = \frac{K'(\lambda_{0,S_+}(\mu_0)) - K(\lambda_1^{**}) + \frac{cp}{v_b}}{v_b(1-\lambda_1^{**})+c} - \frac{p}{v_b}$$

$$\frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0} := \frac{K'(\frac{\mu_1^u(\mu_0)-c}{-v_b})}{v_b} - \frac{p}{v_b}$$

$$\frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0} \Leftrightarrow \frac{K'(\frac{\mu_1^u(\mu_0)-c}{-v_b})}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1-\lambda_1^{**})-c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1-\lambda_1^{**})-c} \qquad (\star)$$

$$\begin{aligned}
& c \leq \tilde{c} = v_g K'^{-1}\left[\frac{K(\lambda_1^{**})}{\lambda_1^{**}}\right] \\
\Leftrightarrow & K(\lambda_1^{**}) \geq \lambda_1^{**} K'(\frac{c}{v_g}) \\
\Rightarrow & -v_b\left[K(\lambda_1^{**}) - \frac{cp}{v_b}\right] \geq (c - v_b\lambda_1^{**})K'(\frac{c}{v_g}) \left(c \leq v_g\lambda_1^{**} \Rightarrow K'(\frac{c}{v_g}) \leq K'(\lambda_1^{**}) = p\right) \\
\Rightarrow & -v_b\left[K(\lambda_1^{**}) - \frac{cp}{v_b}\right] \geq (c - v_b\lambda_1^{**})K'(\lambda_{0,S+}(\mu_0)) \left(\lambda_{0,S+}(\mu_0) < \lambda_{0,S+}(\mu_0') = \frac{c}{v_g}\right) \\
\Leftrightarrow & \frac{K'(\lambda_{0,S+}(\mu_0))}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1-\lambda_1^{**})-c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1-\lambda_1^{**})-c} \\
\overset{Lemma 10}{\Rightarrow} & \frac{K'(\frac{\mu_1^u(\mu_0)-c}{-v_b})}{-v_b} + \frac{K(\lambda_1^{**}) - \frac{cp}{v_b}}{-v_b(1-\lambda_1^{**})-c} \geq \frac{K'(\lambda_{0,S_+}(\mu_0))}{-v_b(1-\lambda_1^{**})-c} \\
\overset{(\star)}{\Leftrightarrow} & \frac{d\Pi_{S_+}(\mu_0)}{d\mu_0} \geq \frac{d\underline{\Pi}_{S_0}(\mu_0)}{d\mu_0}
\end{aligned}$$

One can verify that $\lambda_{0,S_+}(\mu_0') = \frac{c}{v_g}$, $\frac{\mu_1^u(\mu_0')-c}{-v_b} = \lambda_1^{**}$. So, $\Pi_{S_+}(\mu_0') = \underline{\Pi}_{S_0}(\mu_0')$. Therefore, $\underline{\Pi}_{S_0} \geq \Pi_{S_+}(\mu_0)$. $\Pi_{S_0}(\mu_0) \geq \underline{\Pi}_{S_0} \Rightarrow \Pi_{S_0}(\mu_0) \geq \Pi_{S_+}(\mu_0)$.

Now we show that the $S_0$ strategy is feasible whenever the $S_+$ strategy is feasible, which concludes the proof of Proposition 12.

**Lemma 11.** *Suppose $c < v_g\lambda_1^{**}$. For any $\mu_0$ such that $S_+$ strategy is feasible, $S_0$ strategy is also feasible.*

*Proof.* Suppose there exists $\mu_0$ such that $S_+$ strategy is feasible while $S_0$ strategy is not feasible. Then, $2c - v_b\lambda_1^{**} < \frac{2v_g-c}{(v_g)^2}$ and $\frac{(2-\lambda_1^{**})c}{v_g(1-\lambda_1^{**})+c} < \frac{2v_g-c}{(v_g)^2}$, which is equivalent to $\lambda_1^{**} < \frac{c^2}{v_b(v_g)^2} + \frac{2c}{v_g}$ and $c > v_g\lambda_1^{**}$, which is not possible as we assumed that $c < v_g\lambda_1^{**}$. $\square$

$\square$

Proposition 12 tells us that we could limit our attention to $S_0$ strategy when $c \leq \tilde{c}$. Proposition 11 has characterized the optimal $S_0$ strategy. One can see from the characterization of the optimal strategy that the sender's payoff is strictly increasing in the prior.

$\square$

**Proposition 13.** ***(comparative statics wrt the prior when the search cost is intermediate)*** *When the search cost is intermediate, $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$, and the sender provides information in both periods. There exists $\mu_{2,+} \in [\frac{2v_g-c}{(v_g)^2}c, \widehat{\mu}_0)$ and $\mu_{2,0} \in [\frac{2v_g-c}{(v_g)^2}c, \mu_{2,+}]$. The arrival rate of good news in the first period, $\lambda_0^*$, remains the same when $\mu_0 < \mu_{2,0}$ and is strictly increasing in the prior when $\mu_0 > \mu_{2,+}$. The arrival rate of good news in the second period, $\lambda_1^*$, is strictly increasing in the prior when $\mu_0 < \mu_{2,0}$ and remains the same when $\mu_0 > \mu_{2,+}$. The belief after observing good news in the second period, $\bar{\mu}_1^*$, is strictly decreasing in the prior when $\mu_0 < \mu_{2,0}$ or $\mu_0 > \mu_{2,+}$. Good news always fully reveals the state in the first period, $\bar{\mu}_0^* \equiv 1$.*

When the expected receiver surplus in the second period is zero ($S_0$ strategy), the minimum amount of information for the receiver to search in the first period is already too high. Under the $S_+$ strategy, the receiver anticipates the sender to provide favorable information in the second period, which relaxes the first-period participation constraint. Therefore, the receiver is willing to search even if the sender provides less information in the first period. This benefits the sender. However, the $S_+$ strategy has the disadvantage of inducing higher expected receiver search costs.

When the prior is low, the disadvantage of the $S_+$ strategy dominates the advantage. The sender prefers the $S_0$ strategy to the $S_+$ strategy. She faces an information over-provision issue in the first period and an under-provision issue in the second period. In the first period, she provides the minimum amount of information for the receiver to search. So, the arrival rate of good news in the first period, $\lambda_0^*$, does not depend on the prior. In the second period, she provides the maximum amount of information. More frequent good signals are feasible when the prior is higher. Even if the receiver becomes less certain about the state being good after observing good news, he would still search as long as the likelihood of receiving good news and earning a strictly positive surplus increases. In equilibrium, the sender trades off the precision of good news for the frequency as the prior increases.

When the prior is high, the advantage of the $S_+$ strategy dominates the disadvantage. The sender prefers the $S_+$ strategy to the $S_0$ strategy. She faces an information under-provision issue in the first period and no distortion in the second period. In the first period, she provides the maximum amount of information feasible, which is strictly increasing in the prior. In the second period, she provides the optimal amount of information, which does not depend on the prior. When the prior increases, the participation constraint in the first period is relaxed. To persuade the receiver to search in the first period, the sender could provide information less favorable to the receiver in the second period. She trades off the precision of good news in the second period for the frequency of good news in the first period. So, the belief after observing good news, $\bar{\mu}_1^*$, is strictly decreasing in the prior. The optimal strategy is discontinuous when the sender switches the types of strategy (as illustrated in Figure 6).

*Proof of Proposition 13.* We compare the sender surplus between the solution to $(P_{2S_+})$ and the solution to $(P_{2S_0})$. The following result shows that the optimal $S_+$ strategy generates a strictly higher sender surplus than the optimal $S_0$ strategy when the prior is high.

**Lemma 12.** *If $v_g\lambda_0^{**} < c < v_g\lambda_1^{**}$, the receiver gets strictly positive surplus in the second period when $\mu_0$ is close to $\widehat{\mu_0}$.*

*Proof.* According to Proposition 10, the sender achieves the benchmark payoff $-K(\lambda_0^{**}) + p\lambda_0^{**} + (1-\lambda_0^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$ when $\mu_0 \to \widehat{\mu_0}$ by $S_+$ strategy. According to Lemma 8, the sender surplus of $S_0$ strategy is $\leq -K(\frac{c}{v_g}) + \frac{cp}{v_g} + (1-\frac{c}{v_g})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] < -K(\lambda_0^{**}) + p\lambda_0^{**} + (1-\lambda_0^{**})\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right]$ as $\lambda_0^{**} < \frac{c}{v_g} = \lambda_0$. The difference of the benchmark sender surplus and the sender surplus of $S_0$ strategy is larger than a strictly positve constant. So, when $\mu_0 \to \widehat{\mu_0}$, $S_+$ strategy gives the sender strictly higher payoff. $v_g\lambda_0^{**} < c \Leftrightarrow \widehat{\mu_1}(\widehat{\mu_0}) > c - v_b\lambda_1^{**} \Rightarrow$ the receiver gets strictly positive surplus from $S_+$ strategy when $\mu_0 \to \widehat{\mu_0}$. Thus, the receiver gets strictly positive surplus in the second period in equilibrium when $\mu_0 \to \widehat{\mu_0}$. Continuity of the optimal strategy and sender surplus then implies that there exists a neighborhood $[\widehat{\mu_0} - \delta, \widehat{\mu_0}]$ for some $\delta > 0$ such that the receiver gets strictly positive surplus in the second period in equilibrium when $\mu_0 \in [\widehat{\mu_0} - \delta, \widehat{\mu_0}]$. □

When $\frac{c}{v_g} \le \mu_0 < \mu_{1,2}$, the sender provides information in one period. When $\mu_{1,2} \le \mu_0 < \widehat{\mu_0}$, the sender provides information in both periods. Proposition 10, and Lemma 8 imply the explicit form of the optimal strategy. When $\mu_{1,2} \le \mu_0 < \mu_{2,+}$, $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{c}{v_g}, 1)$, $(\lambda_1^*, \bar{\mu}_1^*) = (\frac{v_g(\mu_0-c)-c(1-c)}{p(v_g-c)}, \frac{[v_g\mu_0-c]p}{v_g(\mu_0-c)-c(1-c)})$, $\mu_1^* = \frac{v_g\mu_0-c}{v_g-c} < c - v_b\lambda_1^{**}$. The receiver gets zero surplus at each period. When $\mu_{2,+} \le \mu_0 < \widehat{\mu_0}$, $(\lambda_0^*, \bar{\mu}_0^*) = (\frac{\mu_0-2c+v_b\lambda_1^{**}}{-v_b(1-\lambda_1^{**})-c}, 1)$, $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}})$, $\mu_1^* = \widehat{\mu_1}(\mu_0) > c - v_b\lambda_1^{**}$, where $\widehat{\mu_1}(\mu_0) = \frac{2c-v_b\lambda_1^{**}-(1+c-v_b\lambda_1^{**}+v_b)\mu_0}{c-v_b-\mu_0}$. The receiver gets strictly positive surplus in the second period and zero total surplus. $\square$

*Proof of Propostion 5.* When the search cost is high, $v_g\lambda_1^{**} \le c < \widehat{c}$, the optimal strategy of the sender is $(\lambda_t^*, \bar{\mu}_t{}^*) = (\frac{c}{v_g}, 1), t = 0, 1$ according to Proposition 4, which does not depend on $\eta$.

When the search cost is low, $c < \tilde{c} < v_g\lambda_0^{**}$. The boundary solution does not depend on $\eta$. Consider the interior solution to $(P_{2S_0})$, we have: $\eta\tilde{K}(\frac{\mu_1-c}{-v_b}) + \frac{p-\mu_1}{p}\eta\tilde{K}'(\frac{\mu_1-c}{-v_b}) - \eta\tilde{K}'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + c = 0$. The LHS is strictly increasing in $\mu_1$ and strictly decreasing in $\eta$. So, $\eta \uparrow \Rightarrow \mu_1^* \uparrow \Rightarrow \lambda_1^* = \frac{\mu_1^*-c}{p} \uparrow, \lambda_0^* = \frac{\mu_0-\mu_1^*-c}{p-\mu_1^*} \downarrow$. $\square$

**Proposition 14.** ***(efficient strategy of the subgame)*** *At the second period, the social planner does not provide information when $\mu_1 < \mu_{0,1}$. When $\mu_1 \ge \mu_{0,1}$, the efficient signal fully reveals the state when good news arrives, $\bar{\mu}_{1,e} = 1$; the arrival rate of good news, $\lambda_{1,e}$, depends on the search cost:*

1. *if $c \ge v_g\lambda_1^{**}$, then there exists a unique $\widehat{\widehat{c}} \in (v_g\lambda_1^{**}, \mu_1 v_g]$ such that the social planner does not provide information if $c > \widehat{\widehat{c}}$ and $\lambda_{1,e} = \frac{c}{v_g} \vee (\tilde{\lambda_1} \wedge \mu_1)$ if $c \le \widehat{\widehat{c}}$.*

2. *if $c \in [\mu_1 + v_b\lambda_1^{**}, v_g\lambda_1^{**})$, then $\lambda_{1,e} = \mu_1$.*

3. *if $c < \mu_1 + v_b\lambda_1^{**} \wedge v_g\lambda_1^{**}$, then $\lambda_{1,e} = \tilde{\lambda_1} \wedge \mu_1$.*

*Proof of Proposition 14.* We first introduce a benchmark problem, in which the receiver is forced to participate and the social planner could generate any signal that fully reveals the state when good news arrives. The social planner chooses the information structure to maximize total welfare. We will use the solution throughout the remaining section.

$$\max_{\lambda_1} -K(\lambda_1) + \lambda_1 \tag{$E_b$}$$

**Lemma 13.** *The optimal solution to $(E_b)$ exists and is unique. Denote it by $\tilde{\lambda_1}$. The objective function under $\tilde{\lambda_1}$ is strictly positive. $\tilde{\lambda_1}$ does not depend on the search cost $c$ and $\tilde{\lambda_1} > \lambda_1^{**}$, the solution to the payoff-maximizing benchmark problem.*

*Proof.* All the results follow from the same argument as the proof of Lemma 1 except $\tilde{\lambda_1} > \lambda_1^{**}$. The F.O.C.'s imply $K'(\tilde{\lambda_1}) = 1 > p = K'(\lambda_1^{**}) \Rightarrow \tilde{\lambda_1} > \lambda_1^{**}$. □

As the social planner wants to maximize the total welfare, she always wants to make the precision of the signal, $\bar{\mu_1}$, as high as possible (subject to the feasibility constraint) given $\mu_1$ and $\lambda_1$, to increase receiver surplus while holding sender surplus fixed. So, if the social planner provides information, then $\bar{\mu_1} = \frac{\mu_1}{\lambda_1} \wedge 1$.

1. if $c \geq v_g \lambda_1^{**}$ (*i.e.* $\lambda_1^{**} \leq \frac{c}{v_g}$), then $TS = \begin{cases} -K(\lambda_1) + \mu_1 - c, \ if \ \lambda_1 \geq \mu_1 \\ -K(\lambda_1) + \lambda_1 - c, \ if \ \lambda_1 < \mu_1 \end{cases} \Rightarrow \lambda_{1,e} = \frac{c}{v_g} \vee (\tilde{\lambda_1} \wedge \mu_1)$ when the social planner provides information. The social planner woiuld provide information if the total surplus is non-negative. Similar to the proof of Proposition 2, one can show that there exists a unique $\widehat{\widehat{c}} \in (v_g \lambda_1^{**}, \mu_1 v_g]$ such that the social planner does not provide information iff $c > \widehat{\widehat{c}}$. Moreover, if $\widehat{\widehat{c}} \geq (1-p)\tilde{\lambda_1}$, $\lambda_{1,e} = \frac{c}{1-p} \Rightarrow \bar{\mu}_{1,e} = 1 \Rightarrow TS = -K(\lambda_{1,e}) + p\lambda_{1,e} \Rightarrow \widehat{\widehat{c}} = \widehat{c}$.

2. if $c \in [\mu_1 - p\lambda_1^{**}, v_g \lambda_1^{**})$ (*i.e.* $\lambda_1^{**} \geq \frac{\mu_1 - c}{-v_b} > \frac{c}{1-p}$), then by the previous argument, $\lambda_{1,e} = \frac{c}{1-p} \vee (\tilde{\lambda_1} \wedge \mu_1) = \mu_1$.

3. if $c < \mu_1 - p\lambda_1^{**} \wedge (1-p)\lambda_1^{**}$ (*i.e.* $\lambda_1^{**} \in \left(\frac{c}{1-p}, \frac{\mu_1 - c}{-v_b}\right)$), then the total welfare $TS = \begin{cases} -K(\lambda_1) + \mu_1 - c, \ if \ \lambda_1 \geq \mu_1 \\ -K(\lambda_1) + \lambda_1 - c, \ if \ \lambda_1 < \mu_1 \end{cases} \Rightarrow \lambda_{1,e} = \tilde{\lambda_1} \wedge \mu_1$.

□

*Proof of Proposition 6.* When$(1-p)\tilde{\lambda_1} \leq c \leq \widehat{\widehat{c}}$, the constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0 \bar{\mu_0} - c + (1-\lambda_0)\left[-K(\frac{c}{1-p}) + \frac{cp}{1-p}\right] \qquad (E_{2H})$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \geq \frac{c}{1-p}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to $(E_{2H})$ is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{c}{1-p}, \frac{c}{1-p})$. Therefore, there is no information distortion.

When $v_g\lambda_1^{**} \le c < (1-p)\tilde{\lambda}_1 \wedge \widehat{\widehat{c}}$, the constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0\bar{\mu}_0 - c + (1-\lambda_0)\left[-K(\tilde{\lambda}_1 \wedge \mu_1) + \tilde{\lambda}_1 \wedge \mu_1 - c\right] \tag{$E_{2I}$}$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \ge \frac{c}{1-p}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to $(E_{2I})$ is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{c}{1-p}, \frac{c}{1-p})$. Therefore, there is no information distortion.

When $c < v_g\lambda_1^{**}$, one can see that we could restrict $\mu_1$ to be less than or equal to $\tilde{\lambda}_1$ without loss of generality. The constrained program of the social planner is:

$$\max -K(\lambda_0) + \lambda_0\bar{\mu}_0 - c + (1-\lambda_0)\left[-K(\mu_1) + \mu_1 - c\right] \tag{$E_{2S_0}$}$$
$$\text{s.t. } (IR_0), (F_0), \mu_1 \ge \frac{c}{1-p}$$

Using similar methods of finding the optimal sender strategy in the main text, one can show that the solution to $(E_{2S_0})$ is $(\lambda_{0,e}, \lambda_{1,e}) = (\frac{\mu_0 - \frac{c}{1-p} - c}{p - \frac{c}{1-p}}, \frac{c}{1-p})$. Therefore, the sender provides too much information (relative to the efficient solution) in the second period and too little information in the first period. $\square$

*Proof of Proposition 7.* One can see that if the expected surplus of the receiver in the second period is 0 in the optimal solution to $(P_{dc})$, then $(\lambda_1^*, \bar{\mu}_1^*)$ solves $(P_1)$. Otherwise, the sender can strictly increase the payoff by using the same $(\lambda_0^*, \bar{\mu}_0^*)$ and replacing $(\lambda_1^*, \bar{\mu}_1^*)$ by the optimal solution to $(P_1)$, holding the same $\mu_1$. Hence, if dynamic commitment power strictly increases the sender surplus, the solution to $(P_{dc})$ must satisfy: $\mathbb{E}[\text{receiver surplus at } t = 1] = (1-\lambda_0)[\lambda_1(\bar{\mu}_1 + v_b) - c] > 0$. Denote the optimal sender surplus when the sender does not have dynamic commitment power by $\Pi_{wo}$. The corresponding optimal strategy of the sender is $(\lambda_t^*, \bar{\mu}_t{}^*) = (\frac{c}{v_g}, 1)$ according to Proposition 4. Consider the following strategy: $(\lambda_0, \bar{\mu}_0, \lambda_1, \bar{\mu}_1) = (\frac{c-\delta v_g}{(1-\delta)v_g}, 1, \frac{c}{v_g} + \delta, 1)$. Denote the corresponding sender surplus by $\Pi(\delta)$. One can verify that it is feasible when $\delta > 0$ is small and the sender has dynamic commitment power, and leads to a payoff no larger than the optimal sender surplus with dynamic commitment (denote it by $\Pi_w$).

$$
\begin{aligned}
\Pi(\delta) =& -K\left[\frac{c-\delta v_g}{(1-\delta)v_g}\right] + p\frac{c-\delta v_g}{(1-\delta)v_g} + \left[1-\frac{c-\delta v_g}{(1-\delta)v_g}\right]\left[-K(\frac{c}{v_g}+\delta)+\frac{cp}{v_g}+\delta p\right] \\
\frac{d\Pi(\delta)}{d\delta} =& \frac{v_g-c}{(1-\delta)^2 v_g}\left[I_1(\delta)+I_2(\delta)\right] \\
\text{, where } I_1(\delta) =& K'\left[\frac{c-\delta v_g}{(1-\delta)v_g}\right] - (1-\delta)K'(\frac{cp}{v_g}+\delta) \\
I_2(\delta) =& \frac{cp}{v_g} - K(\frac{c}{v_g}+\delta)
\end{aligned}
$$

$I_1(\delta) \to 0,\ I_2(\delta) \to \frac{cp}{v_g} - K(\frac{c}{v_g}) > 0$ as $\delta \to 0$. Therefore, $\exists \widehat{\delta} > 0\ s.t.\ \frac{d\Pi(\delta)}{d\delta} > 0,\ \forall \delta \in (0, \widehat{\delta}]$. Since $\Pi(\delta)$ is continuous in $\delta$ and $\Pi(0) = \Pi_{wo}$, we have $\Pi_w \geq \Pi(\delta) > \Pi_{wo},\ \forall \delta \in (0, \widehat{\delta}]$.

We now show that the benefit of dynamic commitment power vanishes as the search cost approaches zero in several steps. Denote the optimal sender surplus when the sender has (does not have) dynamic commitment power and the search cost is c by $\Pi_{w,c}$ ($\Pi_{wo,c}$). Note that $\Pi_{w,c} \geq \Pi_{wo,c}, \forall c \geq 0$.

**Lemma 14.** *When the search cost is zero, the optimal sender surpluses with and without dynamic commitment power are the same,* $\Pi_{w,0} = \Pi_{wo,0}$.

*Proof.* Suppose the sender provides information in both periods. When the search cost is 0 and the sender uses one-shot experiments, one can see that $(IR_0)$ and $(IR_1)$ are equivalent to $\bar{\mu}_0 \geq -v_b$ and $\bar{\mu}_1 \geq -v_b$. Therefore, even if the receiver obtains strictly positive surplus in the second period, the first-period participation constraint is not relaxed. Hence, the second-period strategy of the sender maximizes her payoff of the subgame in the solution to $(P_{dc})$, and thus satisfies the constraints of $(P_2)$. To show that dynamic commitment power does not improve the sender's payoff when $c = 0$, we just need to show that iterative signals are not optimal when the sender has dynamic commitment power.[10] When $c = 0 < v_g\lambda_1^{**}$ and $\mu_1 > c - v_b\lambda_1^{**} = -v_b\lambda_1^{**}$, the optimal strategy with and without dynamic commitment power coincides, as $(\lambda_1^*, \bar{\mu}_1^*) = (\lambda_1^{**}, \frac{\mu_1}{\lambda_1^{**}} \wedge 1)$. In the proof of Proposition 1, we have shown that iterative signals are not optimal. So, we just need to show that iterative signals are not optimal when $\mu_1 \leq -v_b\lambda_1^{**}$.

If $\mu_0 \geq -v_b\lambda_1^{**}$, then $\mu_1 = \bar{\mu}_0 > \mu_0 \geq -v_b\lambda_1^{**}$. So, iterative signals are not optimal.

[10] We have shown in Proposition 1 that iterative signals are not optimal when the sender does not have dynamic commitment power.

If $\mu_0 < -v_b\lambda_1^{**}$, we have:

$$\begin{aligned}\Pi_1(\mu_0) =& -K(\frac{\mu_0}{-v_b}) + \frac{\mu_0 p}{-v_b}\\ \Pi_{iter}(\mu_0) =& \max_{\lambda_0,\mu_1} -K(\lambda_0) + \lambda_0\left[-K(\frac{\mu_1}{-v_b}) + \frac{\mu_1 p}{-v_b}\right]\\ &\text{s.t. } (IR_{0,iter}), (IR_{1,iter}), (F_0), (F_1), \mu_1 = \bar{\mu_0}\end{aligned}$$

Denote the optimal soution when the sender uses iterative signals by $(\tilde{\lambda_0}, \tilde{\mu_1})$. $\Pi_{iter}(\mu_0) \leq \Pi_1(\mu_0) \Leftrightarrow -K(\tilde{\lambda_0}) + \tilde{\lambda_0}\left[-K(\frac{\tilde{\mu_1}}{-v_b}) + \frac{\tilde{\mu_1} p}{-v_b}\right] \leq -K(\frac{\mu_0}{-v_b}) + \frac{\mu_0 p}{-v_b}$. To show that iterative signals are not optimal, it suffices to show that $-\tilde{\lambda_0}K(\frac{\tilde{\mu_1}}{-v_b}) + \frac{\tilde{\lambda_0}\tilde{\mu_1}p}{-v_b} \leq -K(\frac{\mu_0}{-v_b}) + \frac{\mu_0 p}{-v_b}$. Strict convexity of $K(\cdot) \Rightarrow \tilde{\lambda_0}K(\frac{\tilde{\mu_1}}{-v_b}) = \tilde{\lambda_0}K(\frac{\tilde{\mu_1}}{-v_b}) + (1-\tilde{\lambda_0})K(0) \geq K(\frac{\tilde{\lambda_0}\tilde{\mu_1}}{-v_b})$. Thus, $-\tilde{\lambda_0}K(\frac{\tilde{\mu_1}}{-v_b}) \leq -K(\frac{\tilde{\lambda_0}\tilde{\mu_1}}{-v_b})$. It suffices to show that $-K(\frac{\tilde{\lambda_0}\tilde{\mu_1}}{-v_b}) + \frac{\tilde{\lambda_0}\tilde{\mu_1}p}{-v_b} \leq -K(\frac{\mu_0}{-v_b}) + \frac{\mu_0 p}{-v_b}$, which hold because $(F_0) \Rightarrow \tilde{\lambda_0}\tilde{\mu_1} \leq \mu_0$ and we know from the F.O.C. that $-K(\lambda) + p\lambda$ is strictly increasing in $\lambda$ when $\lambda < \lambda_1^{**}$ (here, $\lambda \leq \frac{\mu_0}{-v_b} < \lambda_1^{**}$). $\square$

We now make the following observation. Since $\Pi_{w,c} \geq \Pi_{wo,c}, \forall c \geq 0$ and $\Pi_{w,0} = \Pi_{wo,0}$, we must have $\Pi_{wo,c} \to \Pi_{w,c}$ as $c \to 0$ if $\Pi_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$.[11] The next result confirms that it is indeed the case and thus finishes the proof.

**Lemma 15.** $\Pi_{wo,c} \to \Pi_{wo,0}$ *as* $c \to 0$.

*Proof.* Proposition 12 shows that the optimal strategy is $S_0$ strategy when the search cost is low. So, according to Lemma 6 and 7, for any $c$ small enough, the sender's problem is:

$$\Pi_{wo,c} = \max_{\mu_1} -K(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}) + p\,\frac{\mu_0-\mu_1-c}{-v_b-\mu_1} + (1-\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})\left[-K(\frac{\mu_1-c}{-v_b}) + \frac{(\mu_1-c)p}{-v_b}\right] \tag{$P_{2S_0}'''$}$$

$$\text{s.t. } \mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right]$$

Denote the solution when the search cost is 0 by $\mu_{1,0}^*$. Define $\mu_{1,c}$ to be the closest value to $\mu_{1,0}^*$ among $\left[\frac{c}{v_g}, \frac{v_g\mu_0-c}{v_g-c}\right]$ and denote the corresponding sender surplus by $\underline{\Pi}_{wo,c}$. One can see that $\underline{\Pi}_{wo,c} \leq \Pi_{wo,c}$. Since $\mu_{1,0}^* \in \left[0, \frac{v_g\mu_0}{v_g}\right]$, we have $\mu_{1,c} \to \mu_{1,0}^*$ as $c \to 0$. Therefore, $\underline{\Pi}_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$. Since $\underline{\Pi}_{wo,c} \leq \Pi_{wo,c} \leq \Pi_{wo,0}$, we also have $\Pi_{wo,c} \to \Pi_{wo,0}$ as $c \to 0$. $\square$

$\square$

[11] One can easily show this observation formally by the triangle inequality.

*Proof of Proposition 8.* When $c \geq \widehat{c}$, the sender does not provide information for any $\delta$.

(1) $v_g \lambda_1^{**} \leq c < \widehat{c}$

By the same argument as the proof of Proposition 4, one can see that the optimal strategy of the sender does not depend on $\delta$. For low prior, $\mu_0 \in [\frac{c}{v_g}, \frac{2v_g - c}{(v_g)^2}c)$, the sender provides information in one period, $(\lambda^*, \bar{\mu}^*) = (\frac{c}{v_g}, 1)$. For high prior, $\mu_0 \geq \frac{2v_g - c}{(v_g)^2}c$, the sender provides information in both periods, $(\lambda_t^*, \bar{\mu}_t^*) = (\frac{c}{v_g}, 1), t = 0, 1$.

(2) $c \leq \tilde{c}$

The sender's problem can be divided into 2 cases:

$$\begin{aligned} &\max -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\lambda_1^{**}) + p\lambda_1^{**}\right] && (P^{\delta}_{2S_+}) \\ &\text{s.t. } (IR_{0,\delta}), (F_0), \mu_1 \in [c - v_b\lambda_1^{**}, \lambda_1^{**}] \end{aligned}$$

$$\begin{aligned} &\max -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] && (P^{\delta}_{2S_0}) \\ &\text{s.t. } (IR_{0,\delta}), (F_0), \mu_1 \in [\frac{c}{v_g}, c - v_b\lambda_1^{**}] \end{aligned}$$

Consider the solution to $(P^{\delta}_{2S_0})$. By the same argument as the proof of Lemma 6, $(P^{\delta}_{2S_0})$ is equivalent to

$$\begin{aligned} &\max -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] && (P^{\delta'}_{2S_0}) \\ &\text{s.t. } \lambda_0 \in \left[\frac{\mu_0 - \mu_1}{1 - \mu_1}, \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}\right] \\ &\qquad \mu_1 \in \left[\frac{c}{v_g}, \frac{v_g\mu_0 - c}{v_g - c}\right] \end{aligned}$$

By arguments similar to the proof of Lemma 7, $\lambda_0 \leq \frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1}$ is binding for $\mu_0 < \widehat{\mu_0}$. Suppose that $\mu_1$'s constraints are not binding.

The Lagrangian is $\mathcal{L} = -K(\lambda_0) + p\lambda_0 + \delta(1-\lambda_0)\left[-K(\frac{\mu_1 - c}{-v_b}) + \frac{(\mu_1 - c)p}{-v_b}\right] + \eta\left(\frac{\mu_0 - \mu_1 - c}{-v_b - \mu_1} - \lambda_0\right)$ *s.t.* $\eta \geq$

$0, \eta\left(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1}-\lambda_0\right)=0$. F.O.C. $\Rightarrow$

$$\begin{aligned}
\eta &= \delta(p-\mu_1)-\delta\frac{p-\mu_1}{p}K'(\frac{\mu_1-c}{-v_b}) \\
&= -K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})+p+\delta\left[K(\frac{\mu_1-c}{-v_b})-\mu_1+c\right] \\
\Rightarrow &\delta\left[K(\frac{\mu_1-c}{-v_b})+\frac{p-\mu_1}{p}K'(\frac{\mu_1-c}{-v_b})-p+c\right]-K'(\frac{\mu_0-\mu_1-c}{-v_b-\mu_1})+p=0 \qquad (*_\delta)
\end{aligned}$$

The sum of the first two terms of the LHS of $(*_\delta)$ is strictly increasing in $\delta$, and the LHS of $(*_\delta)$ is strictly increasing in $\mu_1$. So, the optimal $\lambda_0$ is strictly decreasing in $\delta$; the optimal $\mu_1$ and $\lambda_1$ are strictly increasing in $\delta$. When one of $\mu_1$'s constraints is binding, $\lambda_0$ and $\lambda_1$ does not depend on $\delta$.

We finish the proof of this case by showing that $S_0$ strategy dominates $S_+$ strategy. Denote the sender surplus of the optimal $S_0$ ($S_+$) strategy when the discount factor is $\delta\in(0,1)$ by $\Pi_{S_0}(\delta)$ ($\Pi_{S_+}(\delta)$) and denote the corresponding optimal $S_0$ ($S_+$) strategy at time t by $\lambda_{t,S_0}(\delta)$ ($\lambda_{t,S_+}(\delta)$). We have:

$$\begin{aligned}
\Pi_{S_0}(\delta) =& -K(\lambda_{0,S_0}(\delta))+p\lambda_{0,S_0}(\delta)+\delta(1-\lambda_{0,S_0}(\delta))\left[-K(\lambda_{1,S_0}(\delta))+p\lambda_{1,S_0}(\delta)\right] \\
\geq& -K(\lambda_{0,S_0}(1))+p\lambda_{0,S_0}(1)+\delta(1-\lambda_{0,S_0}(1))\left[-K(\lambda_{1,S_0}(1))+p\lambda_{1,S_0}(1)\right] \\
=& -K(\lambda_{0,S_0}(1))+p\lambda_{0,S_0}(1)+(1-\lambda_{0,S_0}(1))\left[-K(\lambda_{1,S_0}(1))+p\lambda_{1,S_0}(1)\right] \\
&-(1-\delta)(1-\lambda_{0,S_0}(1))\left[-K(\lambda_{1,S_0}(1))+p\lambda_{1,S_0}(1)\right] \\
\overset{(\dagger)}{\geq}& -K(\lambda_{0,S_+}(1))+p\lambda_{0,S_+}(1)+(1-\lambda_{0,S_+}(1))\left[-K(\lambda_1^{**})+p\lambda_1^{**}\right] \\
&-(1-\delta)(1-\lambda_{0,S_+}(1))\left[-K(\lambda_1^{**})+p\lambda_1^{**}\right] \\
=& -K(\lambda_{0,S_+}(1))+p\lambda_{0,S_+}(1)+\delta(1-\lambda_{0,S_+}(1))\left[-K(\lambda_1^{**})+p\lambda_1^{**}\right] \\
\geq& -K(\lambda_{0,S_+}(\delta))+p\lambda_{0,S_+}(\delta)+\delta(1-\lambda_{0,S_+}(\delta))\left[-K(\lambda_1^{**})+p\lambda_1^{**}\right] \\
=& \Pi_{S_+}(\delta)
\end{aligned}$$

, where the inequality $(\dagger)$ holds because Proposition 12 implies that $-K(\lambda_{0,S_0}(1))+p\lambda_{0,S_0}(1)+(1-\lambda_{0,S_0}(1))\left[-K(\lambda_{1,S_0}(1))+p\lambda_{1,S_0}(1)\right]\geq -K(\lambda_{0,S_+}(1))+p\lambda_{0,S_+}(1)+(1-\lambda_{0,S_+}(1))\left[-K(\lambda_1^{**})+p\lambda_1^{**}\right]$ and we have $\lambda_{0,S_0}(1)\geq\lambda_{0,S_+}(1)$, $-K(\lambda_{1,S_0}(1))+p\lambda_{1,S_0}(1)\leq -K(\lambda_1^{**})+p\lambda_1^{**}$.

□